

GRENOBLE SCIENCES

PETIT TRAITÉ D'INTÉGRATION

RIEMANN, LEBESGUE et KURZWEIL-HENSTOCK

Jean-Yves BRIEND



Petit traité d'intégration

Riemann, Lebesgue et Kurzweil-Henstock

Grenoble Sciences

Grenoble Sciences est un centre de conseil, expertise et labellisation de l'enseignement supérieur français. Il expertise les projets scientifiques des auteurs dans une démarche à plusieurs niveaux (référés anonymes, comité de lecture interactif) qui permet la labellisation des meilleurs projets après leur optimisation. Les ouvrages labellisés dans une collection de Grenoble Sciences ou portant la mention « Sélectionné par Grenoble Sciences » *(Selected by Grenoble Sciences)* correspondent à :

- des projets clairement définis sans contrainte de mode ou de programme,
- des qualités scientifiques et pédagogiques certifiées par le mode de sélection (les membres du comité de lecture interactif sont cités au début de l'ouvrage),
- une qualité de réalisation assurée par le centre technique de Grenoble Sciences.

Directeur scientifique de Grenoble Sciences
Jean Bornarel, Professeur émérite à l'Université Joseph Fourier, Grenoble 1

Pour mieux connaître Grenoble Sciences :
https://grenoble-sciences.ujf-grenoble.fr

Pour contacter Grenoble Sciences :
tél : (33) 4 76 51 46 95, e-mail : *grenoble.sciences@ujf-grenoble.fr*

Livres et pap-ebooks

Grenoble Sciences labellise des livres papier (en langue française et en langue anglaise) mais également des ouvrages utilisant d'autres supports. Dans ce contexte, situons le concept de pap-ebook. Celui-ci se compose de deux éléments :

- un **livre papier** qui demeure l'objet central avec toutes les qualités que l'on connaît au livre papier
- un **site web compagnon** qui propose :
  - des éléments permettant de combler les lacunes du lecteur qui ne posséderait pas les prérequis nécessaires à une utilisation optimale de l'ouvrage,
  - des exercices pour s'entraîner,
  - des compléments pour approfondir un thème, trouver des liens sur internet, etc.

Le livre du pap-ebook est autosuffisant et certains lecteurs n'utiliseront pas le site web compagnon. D'autres l'utiliseront et ce, chacun à sa manière. Un livre qui fait partie d'un pap-ebook porte en première de couverture un logo caractéristique et le lecteur trouvera la liste de nos sites compagnons à l'adresse internet suivante :

https://grenoble-sciences.ujf-grenoble.fr/pap-ebook

Grenoble Sciences bénéficie du soutien de la **région Rhône-Alpes** et du **ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche**.
Grenoble Sciences est rattaché à l'**Université Joseph Fourier** de Grenoble.

ISBN 978 2 7598 1241 7

Petit traité d'intégration

Riemann, Lebesgue et Kurzweil-Henstock

Jean-Yves Briend



17, avenue du Hoggar
Parc d'Activité de Courtabœuf - BP 112
91944 Les Ulis Cedex A - France

Petit traité d'intégration
Riemann, Lebesgue et Kurzweil-Henstock

Cet ouvrage, labellisé par Grenoble Sciences, est un des titres du secteur Mathématiques de la collection Grenoble Sciences d'EDP Sciences, qui regroupe des projets originaux et de qualité. Cette collection est dirigée par Jean Bornarel, Professeur émérite à l'Université Joseph Fourier, Grenoble 1.

Comité de lecture :

- Claude Bardos, Professeur émérite à l'Université Paris-Diderot, Paris 7
- Elie Belorizky, Professeur émérite à l'Université Joseph Fourier, Grenoble 1
- Eric Charpentier, Maître de conférences à l'Université de Bordeaux
- Jean-Pierre Demailly, Professeur à l'Université Joseph Fourier, Grenoble 1
- Bruno Demange, Maître de conférences à l'Université Joseph Fourier, Grenoble 1
- Denise Grenier, Maître de conférences à l'Université Joseph Fourier, Grenoble 1

Cet ouvrage a été suivi par Stéphanie Trine. Mise en page : ARCHITEX ; figures : Sylvie Bordage et Anne-Laure Passavant ; illustration de couverture : Alice Giraud, d'après une figure de Emes2k/de.wikipedia.

Autres ouvrages labellisés sur des thèmes proches (chez le même éditeur) :

Analyse numérique et équations différentielles (J.-P. Demailly) • Exercices corrigés d'analyse avec rappels de cours. Tomes I et II (D. Alibert) • Nombres et algèbre (J.-Y. Mérindol) • Mécanique. De la formulation lagrangienne au chaos hamiltonien (C. Gignoux & B. Silvestre-Brac) • Problèmes corrigés de mécanique et résumés de cours. De Lagrange à Hamilton (C. Gignoux & B. Silvestre-Brac) • Méthodes numériques appliquées pour le scientifique et l'ingénieur (J.-P. Grivet) • Introduction aux variétés différentielles (J. Lafontaine) • Description de la symétrie. Des groupes de symétrie aux structures fractales (J. Sivardière) • Symétrie et propriétés physiques. Des principes de Curie aux brisures de symétrie (J. Sivardière) • Approximation hilbertienne. Splines, ondelettes, fractales (M. Attéia & J. Gaches) • Outils mathématiques à l'usage des scientifiques et ingénieurs (E. Belorizky) • Analyse statistique des données expérimentales (K. Protassov) • Mathématiques pour l'étudiant scientifique. Tomes I et II (P.-J. Haug) • Mathématiques pour les sciences de la Vie, de la Nature et de la Santé (J.-P. Bertrandias & F. Bertrandias) • Introduction à la mécanique statistique (E. Belorizky & W. Gorecki) • Mécanique statistique. Exercices et problèmes corrigés (E. Belorizky & W. Gorecki) • Magnétisme : I Fondements, II Matériaux (sous la direction d'E. du Trémolet de Lacheisserie) • La mécanique quantique. Problèmes résolus. Tomes I et II (V.M. Galitski, B.M. Karnakov & V.I. Kogan) • Éléments de Biologie à l'usage d'autres disciplines. De la structure aux fonctions (P. Tracqui & J. Demongeot) • Minimum Competence in Scientific English (S. Blattes, V. Jans & J. Upjohn)

et d'autres titres sur le site internet
https://grenoble-sciences.ujf-grenoble.fr

Avant-propos

Ce petit livre est né d'une discussion avec Jean-Pierre Demailly au début des années 2000. Chargé d'enseigner l'intégration à des étudiants de deuxième année de ce qui s'appelait alors le DEUG, il m'a fourni l'idée de l'enseigner de manière non conventionnelle en construisant l'intégrale d'une fonction réelle d'après la méthode de Kurzweil et Henstock. Après une première année quelque peu tâtonnante (puis une seconde émaillée d'un mouvement social fort, celui contre le contrat première embauche) ce cours m'a procuré beaucoup de plaisir ainsi, je crois, qu'aux étudiants qui l'ont suivi. Il est en effet passionnant de pouvoir construire l'intégrale de Lebesgue (et même un peu mieux) en n'utilisant que des méthodes et outils simples de l'analyse réelle en une variable.

Ce texte n'a rien d'original et je me suis abreuvé à de nombreuses sources, que j'ai parfois allègrement pillées. Je pourrais citer le texte écrit par Jean-Pierre Demailly soi-même, beaucoup plus rigoureux et complet, mais plus difficile d'accès, ou encore les livres de Robert Bartle ou de Jean Mahwin, ce dernier ouvrage étant malheureusement indisponible depuis longtemps.

Le style du livre pourra heurter certains lecteurs, en particulier les collègues mathématiciens. Je l'ai voulu relativement informel, et ai pris le parti de ne démontrer certains résultats que partiellement (surtout en ce qui concerne les fonctions de plusieurs variables) voire parfois d'en admettre la démonstration. J'ai volontairement limité le chapitre sur les méthodes de calcul explicite d'intégrales, dont l'intérêt me semble faible si l'on ne se lance pas dans des développements poussés, sur le calcul formel par exemple.

J'ai par contre essayé de motiver le plus possible les théorèmes et surtout les définitions, nos étudiants ne se contentant plus d'ingurgiter sans barguigner les montagnes de victuailles mathématiques dont nous aimerions les gaver. Et c'est tant mieux !

Mes remerciements vont à Jean-Pierre Demailly, à Bernard Coupet qui m'a encouragé à enseigner ainsi l'intégrale, et à tous les collègues qui m'ont, parfois à l'insu de leur plein gré, aidé à rédiger ce texte. Je remercie tout particulièrement Peter Haïssinski, Pierre Liardet, Jean-Paul Mohsen, Christophe Pittet, Stéphane Rigat et Hamish Short pour leur relecture de versions préliminaires du manuscrit.

Je me dois enfin de remercier les lecteurs désignés par les éditions Grenoble Sciences qui ont fait un formidable travail permettant de grandement améliorer le texte, ainsi que toute l'équipe de cette maison d'édition pour la qualité de son écoute au service des auteurs.

Et un grand merci à Léon la Migraine, du CDO.

Jean-Yves Briend

Table des matières

Introduction

Le programme de mathématiques de la Terminale S impose d'aborder l'intégrale par une définition de nature géométrique faisant appel à la notion d'aire, définition dont on admet la validité tout en en donnant une justification heuristique par un procédé d'approximation du domaine à métrer. Le point de départ est donc le suivant :

Définition 1 — *Soit f une fonction continue positive sur l'intervalle compact $[a, b]$. On appelle* intégrale de f sur l'intervalle $[a, b]$ *l'aire* A *de la région plane contenue entre l'axe des abscisses et le graphe de f et on la note*

$$\mathrm{A} = \int_a^b f(x)\,\mathrm{d}x.$$

Cette définition s'étend aux fonctions de signe variable en définissant une aire « algébrique ».

La suite du programme donne ensuite les propriétés élémentaires de l'intégrale, comme la linéarité ou la relation de Chasles. Il reste alors le problème du calcul effectif d'intégrales de fonctions autres que les constantes, et l'on énonce alors le théorème suivant, dont la démonstration est le plus souvent admise :

Théorème 1 — *Si f est une fonction dérivable sur l'intervalle compact $[a, b]$,* et si f' est continue, *alors*

$$\int_a^b f'(x)\,\mathrm{d}x = f(b) - f(a).$$

De plus, toute fonction continue g sur $[a, b]$ y admet une primitive G*, unique à l'addition d'une constante près, donnée par la formule*

$$\mathrm{G}(t) = \int_a^t g(x)\,\mathrm{d}x.$$

Cet énoncé, connu sous le nom de *théorème fondamental du calcul différentiel et intégral*, permet le calcul explicite d'intégrales *via* le calcul de primitives.

C'est muni de ces deux résultats/théorèmes que l'élève aborde alors des problèmes concrets de calcul de primitives ou de résolutions d'équations différentielles simples, comme celle du pendule simple par exemple. Mais dès que l'on veut aborder des mathématiques plus avancées ou des applications du calcul intégral à des phénomènes physiques réels, les deux énoncés ci-dessus sont très largement insuffisants. Il faudrait pouvoir disposer de puissants théorèmes d'approximation permettant, par exemple, de calculer la puissance de signaux très bruités (signal radio parasité par un bruit rose ourlé de blanc) ou des intégrales d'origine arithmétique ayant la douloureuse tendance à présenter des pathologies (fonction fâcheusement oscillante).

La première construction complète d'une théorie de l'intégrale permettant de montrer la validité de la définition 1 et d'en déduire une démonstration du théorème fondamental est due à Bernhard Riemann au milieu du 19e siècle. L'intégrale de Riemann a cependant très vite montré quelques limitations, causées en partie par l'impossibilité d'intégrer des fonctions non bornées ou encore par la mauvaise qualité des théorèmes de convergence qui la concernent (hormis celui concernant le passage à la limite uniforme, ils font tous appel à une hypothèse *a priori* d'intégrabilité de la fonction limite).

C'est pour combler ces lacunes qu'Henri Lebesgue inventa, au début du 20e siècle, l'intégrale qui porte maintenant son nom. Le théorème de convergence dominée de Lebesgue, point d'orgue de cette théorie, fait qu'elle est actuellement l'intégrale universellement utilisée, aussi bien en mathématiques qu'en physique. Elle a cependant quelques défauts. L'un d'eux est que sa construction, même pour les fonctions réelles d'une variable réelle, est difficile, et cela rend déraisonnable sa présentation *ab initio* dans un cours de Licence. Par ailleurs, l'intégrale de Lebesgue ne possède pas d'intégrales semi-convergentes, comme par exemple

$$\int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x.$$

Du coup, les rapports entre la théorie de l'intégrale et le calcul différentiel (illustrés par le théorème 1) sont plutôt complexes.

Dans ce livre, nous allons présenter une théorie de l'intégrale due à Jaroslav Kurzweil et Ralph Henstock et datant de la fin des années 1950. Elle est basée sur une modification qui semble anodine de l'intégrale de Riemann. Elle permet, en un coup et avec des outils simples, de pallier les défauts des intégrales de Riemann et de Lebesgue. Elle est basée sur une idée très simple : on calcule des aires à l'aide de sommes de Riemann, sommes basées sur des subdivisions de l'intervalle *dont on adapte le pas à la fonction à étudier*. Cette démarche, qui peut paraître naïve, permet en fait de démontrer une version très générale du théorème fondamental, d'avoir les théorèmes de convergence de l'intégrale de Lebesgue et enfin de traiter les intégrales impropres semi convergentes, inaccessibles aussi bien à Riemann qu'à Lebesgue. Il y a cependant une limitation à cette approche : tout cela marche très bien pour les fonctions *d'une variable réelle*. Dès que l'on aborde les fonctions de deux variables ou plus, la géométrie fait son apparition et rend l'intégrale de Kurzweil-Henstock moins intéressante.

Dans ce cadre, et plus généralement dans toutes les théories de l'intégration sur des espaces plus compliqués que $\mathbf{R}$, ce sont les idées de l'intégrale de Lebesgue qui font la loi.

Le livre est divisé en quatre parties.

La première, bien adaptée aux étudiants de premier cycle, traite de l'intégration des fonctions d'une variable réelle. L'approche est celle de Kurzweil et Henstock du fait de sa simplicité et de son élégance. Elle permet d'attaquer sereinement le calcul intégral classique au travers du calcul des primitives ou encore l'étude de la convergence des intégrales impropres. Elle permet aussi, comme nous l'avons dit, d'aborder des théorèmes de convergence qui sont ceux, peu ou prou, de l'intégrale de Lebesgue. Ce sont d'ailleurs les théorèmes de convergence de l'intégrale de Lebesgue que nous énoncerons et que nous démontrerons dans la seconde partie du livre, parlant au passage d'ensembles de mesure nulle et mettant en place des outils et des méthodes qui seront ceux utilisés plus tard pour l'intégrale des fonctions de plusieurs variables. Cette deuxième partie s'achève par la théorie de Fourier de décomposition des signaux en signaux élémentaires, où l'intégrale de Lebesgue trouve l'une de ses plus belles applications.

Dans la troisième partie, nous étudierons l'intégrale des fonctions de plusieurs variables. Nous adopterons la définition de Kurzweil et Henstock, étendue aux domaines pluridimensionnels. Nous verrons assez vite qu'il nous faudra, si l'on veut garder une théorie simple et cohérente (c.-à-d. avec des théorèmes de changement de variable et de convergence d'énoncé simple et efficace) nous limiter à l'intégrale de Lebesgue, ce que nous ferons bien entendu. Nous pourrons alors introduire quelques théorèmes propres à l'intégration en plusieurs variables, comme le théorème de Fubini ou encore la formule de Stokes, que nous n'étudierons que dans le plan, où elle joue un rôle primordial dans le développement de l'analyse complexe. Ce sera aussi l'occasion d'introduire les idées de la théorie de la mesure, prenant pour exemple principal la mesure de Lebesgue, idées dont nous donnerons quelques applications comme le théorème ergodique, par exemple.

La quatrième et dernière partie propose de nombreux exercices et problèmes portant sur les divers chapitres du livre ainsi qu'un fascicule de résultats.

Chaque chapitre de ce livre se termine par un appendice, donnant parfois la démonstration d'un théorème précédemment admis ou bien traitant d'un sujet en lien avec le chapitre mais quelque peu hors des sentiers battus. Ces appendices sont souvent d'un niveau de difficulté supérieur à celui du reste du texte et la lectrice pourra ne les aborder que pendant ses moments de loisir, une fois les éléments du chapitre correspondant bien maîtrisés et compris.

Le public visé par ce livre est varié. Les deux premières parties pourront intéresser aussi bien des étudiants en licence de mathématique (de la première à la troisième année) qu'en licence de physique ou en classe préparatoire. La troisième partie sera sans doute plus adaptée aux étudiants de master. Enfin l'ouvrage pourra être utile aux étudiants préparant l'agrégation ainsi, nous l'espérons, qu'à tout ceux que l'enseignement de la théorie de l'intégration intéresse.

This page intentionally left blank

Partie I

Intégration des fonctions d'une variable réelle

This page intentionally left blank

Chapitre 1
Quelques rappels d'analyse

L'objet de ce court chapitre est de rappeler quelques notions d'analyse dont nous aurons besoin pour développer sereinement la théorie de l'intégration. Le lecteur est invité à passer directement au chapitre suivant et à ne consulter celui-ci qu'en cas de besoin. Voici, *grosso modo*, les endroits où il lui sera utile :

– les sections sur les nombres réels et les propriétés de complétude de $\mathbf{R}$ sont d'un contenu basique que l'étudiant est censé maîtriser lorsqu'il s'attaque à la théorie de l'intégration ;

– la section sur la topologie de $\mathbf{R}$, centrée sur la notion de compacité, sera utile surtout à partir de la deuxième partie du livre, traitant des théorèmes de convergence de l'intégrale de Lebesgue et des séries de Fourier ;

– la section sur la continuité uniforme est utile pour aborder l'intégrale au sens de Riemann ;

– la section sur la dérivabilité est censée connue et sera utile dès le chapitre 2 ;

– la dernière section sur les divers modes de convergence des suites et séries de fonction ne sera pleinement utilisée qu'à partir de la seconde partie du livre traitant, justement, des liens entre convergence de suites de fonctions et convergence des suites de nombres obtenues en intégrant ces fonctions.

1.1. Nombres réels

La théorie de l'intégration permet de définir la moyenne d'une quantité dépendant d'un paramètre *continu.* La notion de moyenne est simple et bien connue pour les quantités dépendant d'un paramètre *discret* et l'intégrale est avant tout une généralisation de cette moyenne discrète : il s'agit donc d'un passage typique *du discret au continu.* C'est pourquoi nous commençons ce chapitre par quelques rappels sur le corps des réels et ses liens avec les nombres entiers et rationnels.

Nous noterons $\mathbf{N}$ l'ensemble des entiers positifs, à savoir $\mathbf{N} = \{0, 1, 2, \ldots\}$, et $\mathbf{Z}$ l'anneau des entiers relatifs. Le corps des nombres rationnels sera noté $\mathbf{Q}$. Enfin nous noterons $\mathbf{R}$ l'ensemble des nombres réels. Rappelons que $\mathbf{R}$, muni des deux opérations d'addition et de multiplication est un corps commutatif et que l'on convient des inclusions suivantes :

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R}.$$

L'ensemble des nombres réels est *ordonné* par une relation d'ordre usuelle que nous notons, comme tout le monde, $\leqslant$. C'est une relation d'ordre total qui est compatible avec la structure algébrique. Cette notion d'ordre permet de définir les différents types d'intervalle rencontrés en analyse, plus particulièrement les intervalles *ouverts* de la forme $]a, b[= \{x \in \mathbf{R}, a < x \text{ et } x < b\}$ et les intervalles fermés ou *segments* de la forme $[a, b] = \{x \in \mathbf{R}, a \leqslant x \text{ et } x \leqslant b\}$. Elle nous donne aussi la notion de *valeur absolue,* qui permet de définir la *distance* entre deux nombres réels a, b par $|a - b|$ qui n'est autre, aussi, que la *longueur* de tout intervalle non vide de bornes a, b.

L'analyse mathématique s'attache à déterminer ce qui est grand et ce qui est petit. L'échelle de toute chose étant l'unité, les nombres entiers doivent fournir la mesure de tout nombre, au sens où tout nombre, aussi grand soit-il, doit pouvoir être dépassé par un nombre entier. L'ensemble ordonné $\mathbf{R}$ possède cette caractéristique : il satisfait à la *propriété d'Archimède* : pour tout nombre réel x, il existe un nombre entier n tel que $|x| \leqslant n$. De manière équivalente, pour tout $\varepsilon > 0$ aussi petit soit-il, il existe un entier $n > 0$ tel que $1/n \leqslant \varepsilon$. L'une des conséquences simples de la propriété d'Archimède est que l'ensemble des rationnels est *dense* dans l'ensemble des réels.

Le corps $\mathbf{R}$ est censé nous permettre de repérer des points sur une droite *continue.* Les propriétés énoncées jusqu'ici sont insuffisantes à cela. Ainsi sont-elles toutes satisfaites par $\mathbf{Q}$ soi-même, mais un simple petit dessin avec règle et compas nous dira qu'un objet ayant les prétentions de $\mathbf{R}$ doit contenir $\sqrt{2}$ et tout un tas d'autres nombres irrationnels. Dans la construction au compas, le nombre $\sqrt{2}$ apparaît car un cercle (centré en un sommet d'un carré unité et passant par le sommet opposé) vient *couper* une droite (celle portant l'un des deux côtés adjacent au sommet du carré centre du cercle). On peut donner deux sens au verbe « couper ». Le premier est celui d'une séparation en deux (ou plus) de parties comme lorsque l'on découpe une tarte en parts, par exemple. Le second est celui qui a trait à la rencontre de deux trajectoires ou courbes, comme lorsque l'on dit que deux routes se coupent en un certain carrefour. Dire qu'un objet, ici $\mathbf{R}$, est continu, c'est identifier ces deux sens et dire qu'ils se correspondent intégralement.

Nous appellerons *coupure* de $\mathbf{R}$ toute paire de parties (G, D) de $\mathbf{R}$ telles que :

- $\{\mathrm{G}, \mathrm{D}\}$ forme une *partition* de $\mathbf{R}$ au sens où $\mathrm{G} \cup \mathrm{D} = \mathbf{R}$ et où $\mathrm{G} \cap \mathrm{D} = \emptyset$;
- G est à gauche de D au sens où pour tout $x \in \mathrm{G}$ et tout $y \in \mathrm{D}$ on a $x \leqslant y$.

Si l'on se donne un nombre réel x on peut naturellement définir deux coupures (G, D) et $(\mathrm{G}', \mathrm{D}')$ de $\mathbf{R}$ par

$$\mathrm{G} =]-\infty, x], \ \ \mathrm{D} =]x, +\infty[\quad \text{et} \quad \mathrm{G}' =]-\infty, x[, \ \ \mathrm{D}' = [x, +\infty[.$$

Le corps ordonné **R** est caractérisé par le fait que les deux notions de coupure sont équivalentes :

Propriété des coupures — Une coupure se fait nécessairement en un nombre réel : pour toute coupure (G, D) de **R** il existe un nombre réel et un seul x tel que $G =]-\infty, x]$ ou $D = [x, +\infty[$.

L'existence d'un corps tel que **R** peut être prouvée dans le cadre de la théorie des ensembles, en construisant les choses de manière progressive à partir de l'ensemble **N** des entiers naturels : on construit **Z** par un procédé de symétrisation puis **Q** par la construction de ce que l'on appelle le corps des fractions. Jusque là nous sommes dans le domaine de l'algèbre. Pour construire **R**, on dispose de plusieurs approches, conduisant toutes au même résultat. On peut par exemple définir des coupures de **Q**, qu'on appelle alors des coupures de Dedekind, de la même manière que l'on a défini plus haut des coupures de **R**, et définir **R** comme *l'ensemble des coupures de Dedekind.* À charge alors au mathématicien de définir sur l'objet ainsi construit une addition, une multiplication et un ordre satisfaisants. Une autre approche consiste à rajouter à **Q** toutes les limites possibles de suites de Cauchy d'éléments de **Q** (cf. *infra*), ce qui constitue la construction dite *de complétion.* On pourrait aussi considérer les nombres en écriture décimale et rajouter aux décimaux toutes les écritures illimitées possibles, ou encore s'inspirer de la définition 6 du livre V des Éléments d'Euclide, pour construire les nombres réels *à la* Eudoxe...

1.2. Propriétés de complétude de R

On peut démontrer que **R** est l'unique corps commutatif totalement ordonné, archimédien et satisfaisant à la propriété des coupures. Cette dernière est cependant rarement utilisable telle quelle et l'on utilise souvent les quelques résultats qui suivent.

Théorème 1.1 — *Toute suite croissante et majorée de nombre réels est convergente.*

Théorème 1.2 — *Toute partie non vide et majorée de* **R** *admet une borne supérieure.*

Le premier théorème est une conséquence simple de la propriété des coupures. Le second, dit « propriété de la borne supérieure » est équivalent à celle des coupures. Il est également équivalent à la conjonction de la propriété d'Archimède et de la convergence des suites croissantes et majorées. On déduit facilement de tout cela le *théorème des segments emboîtés* :

Théorème 1.3 — *Soit* (I_n) *une suite de segments emboîtés dont la longueur tend vers* 0 *quand* n *tend vers l'infini. Alors il existe un nombre réel* x *et un seul tel que pour tout* $n \in \mathbf{N}$ *on ait* $x \in I_n$.

DÉMONSTRATION — Posons $\mathrm{I}_n = [a_n, b_n]$. De l'emboîtement des intervalles on déduit que la suite (a_n) est croissante et la suite (b_n) décroissante et que $a_n \leqslant b_n$ pour tout n. La suite (a_n) est clairement majorée, par exemple par b_0 ; elle est donc convergente et nous noterons x sa limite. De même la suite (b_n) est décroissante et minorée et nous appellerons y sa limite. Montrons par l'absurde que $x \leqslant y$. Supposons donc que $y < x$. Du fait que $x = \lim(a_n)$, il existe un entier N tel que pour tout $n \geqslant \mathrm{N}$ on ait $a_n > x - (x-y)/2$. De même du fait que $y = \lim(b_n)$ il existe M tel que pour tout $n \geqslant \mathrm{M}$ on ait $b_n < y + (x-y)/2$. Pour n plus grand que N et M, les nombres a_n sont à droite du milieu du segment $[y, x]$ et les nombres b_n à gauche. En particulier $b_n < a_n$, ce qui est absurde. Ainsi on a bien $x \leqslant y$. En utilisant le même raisonnement et l'hypothèse que la longueur des intervalles tend vers 0, on montre qu'en fait $x = y$ et donc que (a_n) et (b_n) ont une limite commune x qui satisfait, par monotonicité des suites, a $a_n \leqslant x \leqslant b_n$ pour tout n. Mais alors pour tout $n \in \mathbf{N}$ on a $x \in \mathrm{I}_n$, ce qui achève la démonstration. □

Si l'on relit la preuve qui vient d'être faite, on s'aperçoit que l'on a démontré le résultat suivant, dit « théorème des gendarmes » :

Théorème 1.4 — *Soient (a_n) et (b_n) deux suites de nombres réels. On suppose que la suite (a_n) est croissante tandis que la suite (b_n) est décroissante. On suppose de plus que les deux suites sont* adjacentes *:*

$$\text{pour tout } n \in \mathbf{N},\ a_n \leqslant b_n \text{ et } \lim_{n\to+\infty}(a_n - b_n) = 0.$$

Alors les deux suites (a_n) et (b_n) sont convergentes et admettent la même limite.

Rappelons pour finir la notion de suite de Cauchy et la complétude de $\mathbf{R}$:

Définition 1.1 — *Une suite (x_n) de nombres réels est dite de Cauchy si pour tout $\varepsilon > 0$ il existe un entier $\mathrm{N} \geqslant 0$ tel que, pour tous $n, m \geqslant \mathrm{N}$ on ait $|x_n - x_m| < \varepsilon$.*

Il est facile de démontrer qu'une suite qui converge est de Cauchy. On déduit de la propriété de la borne supérieure le théorème suivant, affirmant que la réciproque est vraie :

Théorème 1.5 — *Toute suite de Cauchy de nombres réels admet une limite. On dit que $\mathbf{R}$ est* complet.

Ce théorème est très utile : il permet de démontrer qu'une suite converge *sans connaître* a priori *sa limite.*

1.3. Topologie de R

Nous utiliserons parfois un peu de terminologie provenant de la topologie. Rappelons qu'un ouvert de $\mathbf{R}$ est une partie $\mathrm{U} \subset \mathbf{R}$ telle que pour tout $x \in \mathrm{U}$, il existe un

intervalle ouvert contenant x et contenu dans U. Un fermé est une partie de $\mathbf{R}$ dont le complémentaire est ouvert. Les ouverts sont appréciés pour la raison suivante : autour d'un point d'un ouvert U, on dispose de place pour bouger tout en restant dans U. Les fermés ont, eux, la propriété complémentaire qui dit que si une suite d'éléments d'un fermé F converge, alors la limite est encore dans F.

Si (x_n) est une suite de nombres réels, on dit que $a \in \mathbf{R}$ est *valeur d'adhérence* de la suite (x_n) s'il existe une sous-suite extraite de (x_n) qui admet a pour limite. Dit autrement, pour tout $\varepsilon > 0$, il existe une infinité d'indices n tels que x_n est dans l'intervalle $]a-\varepsilon, a+\varepsilon[$. Le théorème suivant, aussi connu sous le nom de théorème de Bolzano-Weierstraß, est d'une importance cruciale en analyse :

Théorème 1.6 — *Toute suite bornée de nombres réels admet une valeur d'adhérence.*

Nous allons en donner une démonstration car le schéma de preuve, basé sur le théorème des segments emboîtés, sera réutilisé plus loin.

Démonstration — Soit (x_n) une suite bornée de nombres réels. Il existe deux nombres m, M tels que, pour tout $n \in \mathbf{N}$, on ait $m \leqslant x_n \leqslant \mathrm{M}$. Notons I_0 le segment $[m, \mathrm{M}]$ qui contient donc tous les éléments de la suite (x_n). Considérons les deux intervalles moitiés $[m, (m+\mathrm{M})/2]$ et $[(m+\mathrm{M})/2, \mathrm{M}]$. L'un de ces deux intervalles, notons-le I_1, doit contenir une infinité d'éléments de la suite (x_n). Nous pouvons réitérer ce procédé et ainsi définir l'intervalle I_2 qui est l'une des moitiés de I_1 et contient une infinité d'éléments de la suite (x_n). Cela permet de construire une suite (I_n) de segments emboîtés, dont la longueur tend vers 0 et qui tous contiennent une infinité d'éléments de la suite (x_n). Le point commun à tous ces intervalles dont l'existence nous est assurée grâce au théorème des segments emboîtés est clairement une valeur d'adhérence de la suite (x_n). □

Si une suite ne prend qu'un nombre fini de valeurs $\{a, b, c \ldots\}$, elle admet bien évidemment des valeurs d'adhérence et celles-ci sont prises parmi $\{a, b, c \ldots\}$. Il est tentant d'essayer de trouver une caractérisation des parties de $\mathbf{R}$ qui assurent que, si une suite est à valeurs dans cette partie, elle admet nécessairement une valeur d'adhérence dans cette partie. L'idée est que l'on peut quelque peu épaissir les parties finies, ou plutôt qu'une telle partie, si on l'examine avec une précision finie $\varepsilon > 0$, se comporte comme une partie finie. C'est la notion de compacité.

Définition 1.2 — *Soit* X *une partie de* $\mathbf{R}$. *On dit que* X *est compacte si de toute famille d'ouverts non vides* $\mathcal{U}$ *dont l'union contient* X, *on peut extraire une sous-famille finie dont l'union contient* X. *La famille* $\mathcal{U}$ *s'appelle un* recouvrement ouvert *de* X *et la sous-famille que l'on en extrait un* sous-recouvrement.

Le théorème suivant, connu sous le nom de théorème de Borel-Lebesgue ou de Heine-Borel, caractérise les compacts de $\mathbf{R}$.

Théorème 1.7 — *Soit* X *une partie non vide de* **R**. *Les propriétés suivantes sont équivalentes :*

1. *La partie* X *est compacte ;*
2. *La partie* X *est fermée et bornée ;*
3. *Toute suite d'éléments de* X *admet une sous-suite qui converge vers un élément de* X.

Ce théorème s'étend aux parties X *de* $\mathbf{R}^d$ *pour* $d \in \mathbf{N}$ *quelconque, mais est faux si l'on s'intéresse aux parties d'un espace de dimension infinie.*

Nous donnerons dans le troisième chapitre une preuve élémentaire de ce théorème.

1.4. Continuité et continuité uniforme

Continuons ce court chapitre par deux notions importantes. Nous en donnerons une réinterprétation un peu plus loin, en termes de jauges et de subdivisions.

Définition 1.3 — *Une fonction* f *définie sur une partie* X *de* **R** *est dite continue sur* X *si, pour tout* $x \in \mathrm{X}$ *et tout* $\varepsilon > 0$*, il existe un* $\delta > 0$ *(qui dépend de* f, ε et x*) tel que, si* $y \in \mathrm{X}$ *et* $|y - x| < \delta$ *alors* $|f(y) - f(x)| < \varepsilon$.

Dans cette définition, le nombre $\delta > 0$ que l'on trouve peut dépendre du point $x \in \mathrm{X}$ en lequel on estime la continuité de f. Prenons l'exemple de $f(x) = x^2$, définie sur **R**. De $(x+t)^2 = x^2 + 2tx + t^2$ on déduit que si $\delta < \sqrt{|x|^2 + \varepsilon} - |x|$ on a $|y^2 - x^2| < \varepsilon$ dès que $|y - x| < \delta$. Plus x est grand, ε étant fixé, plus nous devrons prendre δ petit. Si, pour une certaine fonction f, on peut prendre un δ indépendant de x, on dit que f est uniformément continue.

Définition 1.4 — *Une fonction* f *définie sur une partie* X *de* **R** *est dite uniformément continue sur* X *si, pour tout* $\varepsilon > 0$*, il existe un nombre réel* $\delta > 0$ *(qui donc ne dépend que de* f *et* ε*) tel que, pour tous* $x, y \in \mathrm{X}$*,* $|x - y| < \delta$ *entraîne* $|f(x) - f(y)| < \varepsilon$.

Ces deux définitions illustrent bien l'importance de l'ordre des quantificateurs dans un énoncé. Nous avons vu avec la fonction x^2 qu'il existe des fonctions continues mais non uniformément continues. Le théorème suivant nous assure cependant que sur les compacts, toute fonction continue est uniformément continue : c'est le théorème de Heine.

Théorème 1.8 — *Soit* f *une fonction continue définie sur une partie* compacte X *de* **R**. *Alors* f *est uniformément continue sur* X. *Cela s'applique en particulier aux fonctions continues sur un* segment *de* **R**.

Ce théorème est très important dans la théorie de Riemann de l'intégration, où l'on découpe des intervalles en petits bouts de taille fixe, ce qui réclame un contrôle uniforme des variations de la fonction. Nous pouvons finir ces rappels par l'énoncé du théorème suivant, dont nous donnerons une démonstration dans l'appendice au chapitre 3.

Théorème 1.9 — *Une fonction continue sur un compact* X *de* **R** *est bornée et atteint ses bornes.*

Remarquons que tout ce que nous venons de dire s'étend au cas où X est une partie de $\mathbf{R}^d$ et f une fonction à d variables définies sur X.

1.5. Dérivabilité et inégalité des accroissements finis

Nous rappelons maintenant les quelques notions et théorèmes fondamentaux relatifs à la dérivée des fonctions.

Définition 1.5 — *Soient* I *un intervalle ouvert de* **R**, x *un élément de* I *et* f *une fonction définie sur* I. *On dit que* f *est dérivable en* x *si les taux d'accroissements*

$$\frac{f(y)-f(x)}{y-x},\ y \neq x,$$

admettent une limite finie quand y *tend vers* x. *On note alors* $f'(x)$ *cette limite qui s'appelle le nombre dérivé de la fonction* f *en* x. *On dit que* f *est dérivable sur* I *si elle est dérivable en tout point de* I. *La fonction* $x \mapsto f'(x)$ *est alors appelée la* dérivée *de la fonction* f.

Il est facile de démontrer qu'une fonction dérivable sur I y est aussi continue et que la réciproque est fausse (nous renvoyons au chapitre d'exercices où sera exhibée une fonction continue et nulle part dérivable, la fonction de Weierstraß). Enfin on peut étendre la notion de dérivabilité, comme pour la continuité, aux fonctions définies sur un intervalle fermé, en définissant les notions de dérivées à droite et à gauche.

On peut réécrire la définition de la dérivabilité de f sur I de la manière suivante : pour tout $\varepsilon > 0$, pour tout $x \in \mathrm{I}$, il existe δ (dépendant donc de x) et un nombre réel $f'(x)$ tels que, pour tout $y \in \mathrm{I}, y \neq x$, on ait

$$\left|\frac{f(y)-f(x)}{y-x} - f'(x)\right| < \varepsilon,$$

ce que l'on peut encore écrire, en autorisant cette fois-ci y à être égal à x :

$$|f(y) - f(x) - (y - x)f'(x)| < \varepsilon|f'(x)|,$$

inégalité qui signifie qu'au voisinage de x la fonction f est proche de la fonction affine $y \mapsto f(x) + (y - x)f'(x)$. Il faut noter que, comme pour la définition de la continuité, le δ que l'on considère dépend du point x. La notion correspondant à l'uniforme continuité est celle de fonction de classe $\mathcal{C}^1$.

Définition 1.6 — *Une fonction f définie sur un intervalle ouvert* I *est dite de classe $\mathcal{C}^1$ si f est dérivable sur* I *et si sa dérivée f' est continue sur* I*. Plus généralement, si $k \geqslant 1$ est un entier, on dit que f est de classe $\mathcal{C}^k$ si f est k fois dérivable sur* I *et si sa dérivée k^{e} notée $f^{(k)}$ est continue sur* I*. Par convention, une fonction est dite de classe $\mathcal{C}^0$ si elle est continue.*

Si l'on note $\mathcal{C}^k(\mathrm{I})$ l'ensemble des fonctions de classe $\mathcal{C}^k$ sur l'intervalle I, alors pour l'addition et la multiplication usuelle des fonctions $\mathcal{C}^k(\mathrm{I})$ forme un anneau unitaire. On a de plus la série d'inclusions $\mathcal{C}^0(\mathrm{I}) \supset \mathcal{C}^1(\mathrm{I}) \supset \cdots \supset \mathcal{C}^k(\mathrm{I}) \supset \cdots$ et l'intersection de tous ces anneaux forme l'anneau des fonctions de classe $\mathcal{C}^\infty$. Pour montrer l'analogie entre les fonctions de classe $\mathcal{C}^1$ et les fonctions uniformément continues, nous pouvons énoncer le théorème suivant.

Théorème 1.10 — *Soient* I *un intervalle compact et f une fonction définie sur* I*. La fonction f est de classe $\mathcal{C}^1$ sur* I *si et seulement si f est uniformément dérivable sur* I*, au sens où il existe une fonction f' définie sur* I *telle que, pour tout $\varepsilon > 0$, il existe un δ tel que, pour tous $x, y \in \mathrm{I}$ on ait*

$$|f(y) - f(x) - (y - x)f'(x)| < \varepsilon|f'(x)|.$$

La dérivabilité a deux conséquences importantes. La première est qu'elle permet un contrôle local de la fonction f au moyen de la théorie des développements limités. Mais elle permet aussi un contrôle *global* des variations de la fonction. C'est le contenu de l'inégalité des accroissements finis.

Théorème 1.11 — *Soit f une fonction définie et continue sur le segment $[a, b]$, telle que f soit dérivable sur $]a, b[$ et que la dérivée de f sur $]a, b[$ soit bornée. Alors pour tous $x, y \in [a, b]$ on a*

$$|f'(x) - f'(y)| \leqslant |x - y| \sup\{|f'(t)|, t \in]a, b[\}.$$

Nous donnerons dans le chapitre 3 une preuve élémentaire de l'inégalité des accroissements finis. Elle est souvent présentée comme conséquence du théorème suivant, connu sous le nom de *formule des accroissements finis.*

Théorème 1.12 — *Soit f une fonction définie et continue sur le segment $[a, b]$ et dérivable sur l'intervalle $]a, b[$. Il existe alors un élément c de $]a, b[$ tel que*

$$f(b) - f(a) = (b - a)f'(c).$$

On peut démontrer la formule des accroissements finis en utilisant le lemme de Rolle, qui est une conséquence simple du fait qu'une fonction continue sur un segment est bornée et atteint ses bornes, et du fait qu'en un maximum ou un minimum, une fonction dérivable est de dérivée nulle.

1.6. Modes de convergence des suites et séries de fonctions

Des pans entiers de l'analyse mathématique consistent en l'étude de l'approximation de fonctions compliquées par des fonctions simples. Nous en verrons un exemple dans le chapitre traitant des séries de Fourier. Il s'agit d'écrire une fonction f définie sur un intervalle I comme limite d'une suite de fonctions (f_n) elles aussi définies sur I et plus simples à appréhender. Il y a plusieurs manières, pour une suite de fonctions, de converger vers une fonction limite f. Nous allons ici rappeler la signification de ces divers modes de convergence. Le plus naïf est celui de la convergence simple.

Définition 1.7 — *Soit (f_n) une suite de fonctions définies sur un intervalle I. On dit que (f_n) converge* simplement *vers une fonction f définie sur I si, pour tout $x \in \mathrm{I}$, on a*

$$\lim_{n \to +\infty} f_n(x) = f(x).$$

On parle aussi de convergence ponctuelle, et on dit que f est la limite simple *de la suite (f_n).*

Explicitons un peu plus cette convergence simple. Soit $\varepsilon > 0$. Pour chaque élément $x \in \mathrm{I}$, il existe un entier N, qui dépend de x et de ε et que nous noterons donc $\mathrm{N}_\varepsilon(x)$, tel que, si $n \geqslant \mathrm{N}_\varepsilon(x)$, on a $|f_n(x) - f(x)| < \varepsilon$. Dans ce mode de convergence on ne dispose d'aucune information sur la dépendance de N_ε en fonction de x. Si $\mathrm{N}_\varepsilon(x)$ peut être majoré, pour tout x, par un nombre N *indépendant de x*, alors on dit que (f_n) converge *uniformément* vers f.

Définition 1.8 — *Soit (f_n) une suite de fonctions définies sur un intervalle I. On dit que (f_n) converge* uniformément *vers une fonction f définie sur I si, pour tout $\varepsilon > 0$, il existe un entier $\mathrm{N} \geqslant 0$ tel que, pour tout $n \geqslant \mathrm{N}$ et* tout *$x \in \mathrm{I}$ on ait*

$$|f_n(x) - f(x)| < \varepsilon.$$

On dit que f est la limite uniforme *de la suite (f_n).*

Les limites simples de fonctions simples, comme par exemple les fonctions continues, peuvent être assez sauvages. La convergence uniforme, elle, fait passer beaucoup de propriétés des fonctions de la suite à la limite. Nous énonçons en ce sens deux théorèmes importants.

Théorème 1.13 — *Si* (f_n) *est une suite de fonctions* continues *sur l'intervalle* I *et si* (f_n) *converge uniformément, alors la limite de la suite est elle aussi continue sur* I.

De même que la continuité, la dérivabilité passe bien, sous certaines hypothèses, aux limites uniformes :

Théorème 1.14 — *Soit* (f_n) *une suite de fonctions de classe* $\mathcal{C}^1$ *définies sur l'intervalle* I. *Si* (f_n) *converge simplement vers une fonction* f *et si, de plus, la suite* (f'_n) *des dérivées converge uniformément vers une fonction* g*, alors la fonction* f *est de classe* $\mathcal{C}^1$ *et sa dérivée est égale à* g.

Pour illustrer le fait que l'on peut approcher des fonctions compliquées par des fonctions simples, énonçons le théorème de Stone-Weierstraß :

Théorème 1.15 — *Soient* $[a, b]$ *un segment de* **R** *et* f *une fonction continue sur* $[a, b]$. *Il existe une suite de polynômes* (P_n) *telle que* (P_n) *converge uniformément vers* f *sur* $[a, b]$.

Comme nous le verrons lorsque nous traiterons des séries de Fourier, il est souvent utile de considérer également des *séries* de fonctions. Rappelons que l'on parle de série lorsqu'étant donné une suite (f_n) (le terme général de la série) on s'intéresse à la suite des sommes partielles

$$\mathrm{S}_n = \sum_{k=0}^{n-1} f_n.$$

Nous pouvons bien évidemment étendre à ce cadre les notions de convergence simple et uniforme : la série de terme général (f_n) (que l'on note parfois Σf_n) converge simplement (resp. uniformément) vers f si la suite des sommes partielles de la série converge simplement (resp. uniformément) vers f. Mais pour les séries de fonctions on dispose de deux autres modes de convergence.

Définition 1.9 — *On dit que la série de fonctions* Σf_n *définie sur* I *converge absolument si, pour tout* $x \in \mathrm{I}$*, la série numérique* $\Sigma f_n(x)$ *converge absolument.*

Rappelons que, du fait que **R** est complet, si une série converge absolument alors elle converge tout court. Ainsi une série de fonctions absolument convergente est-elle simplement convergente.

Définition 1.10 — *On dit que la série de fonctions* Σf_n *définie sur* I *converge normalement si la série numérique de terme général* $\sup_{x\in \mathrm{I}} |f_n(x)|$ *converge.*

Il est clair que si la série converge normalement alors elle converge absolument. Mais, mieux, si elle converge normalement alors elle converge uniformément.

1.7. Appendice : les nombres réels d'Eudoxe

Nous allons dans cette section présenter une construction des nombres réels souvent appelée « des nombres réels d'Eudoxe ». Elle n'est pas, bien évidemment, due à Eudoxe de Cnide, qui vécut au 4e siècle avant notre ère. Cependant elle s'inspire du cinquième livre des Éléments d'Euclide qui, paraît-il, est dû à Eudoxe. Ce livre traite des proportions qui sont des entités de nature géométrique qui ne peuvent, à l'époque, être considérées comme des nombres (penser à la proportion de la diagonale d'un carré à l'un de ses côtés). Le début du livre est consacré aux définitions et l'une d'entre elles définit l'égalité de deux proportions. Dans la traduction de F. Peyrard, on lit, sous le numéro 6, la définition suivante :

> Des grandeurs sont dites être en même raison, la première à la seconde et la troisième à la quatrième, lorsque les équimultiples quelconques de la première et de la troisième, et d'autres équimultiples quelconques de la seconde et de la quatrième sont tels, que les premiers équimultiples surpassent, chacun à chacun, les seconds équimultiples, ou leur sont égaux à la fois, ou plus petits à la fois.

La traduction en langage algébrique moderne est relativement aisée. Tout d'abord le mot « raison » réfère à un *ratio,* c'est-à-dire à un rapport (rappelons qu'il s'agit du livre traitant des proportions). Ensuite lorsqu'Euclide parle d'équimultiples de grandeurs, il entend par là que l'on prend des multiples de ces deux grandeurs par *un même nombre entier*. La définition 6 peut alors se traduire de la manière suivante, si l'on fait l'identification sacrilège des grandeurs et des nombres réels : le rapport a_1/a_2 est égal au rapport a_3/a_4 si et seulement si, pour tous entiers $n, m \in \mathbf{Z}$, on a ($ma_1 > na_2$ et $ma_3 > na_4$) ou ($ma_1 = na_2$ et $ma_3 = na_4$) ou ($ma_1 < na_2$ et $ma_3 < na_4$).

Nous allons simplifier cette définition d'égalité et la rendre plus géométrique : un rapport peut s'interpréter comme la *pente* d'une droite dans le plan. Donnons-nous en effet deux couples de segments de grandeurs respectives a, b et a', b'. Donnons-nous également un point O dans le plan. Du point O, traçons horizontalement le segment de grandeur a puis, de l'extrémité de a, élevons verticalement le segment de grandeur b. Nous atteignons ainsi un point que nous noterons B. Faisons de même avec a' et b', obtenant le point B′. Le théorème de Thalès nous assure alors que a, b et a', b' sont en même raison si et seulement si les points $\mathrm{O}, \mathrm{B}, \mathrm{B}'$ sont alignés. La définition 6 propose de transcrire cette égalité en ne regardant que les équimultiples des segments.

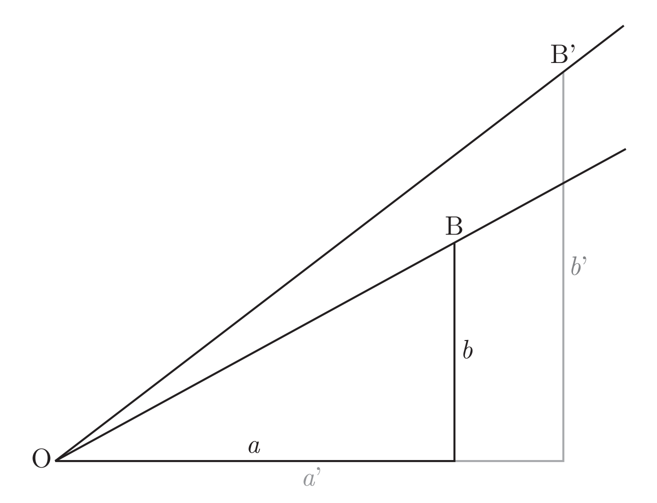


Figure 1.1 – Deux proportions inégales

Traduisons cela géométriquement. Pour plus de simplicité nous allons adopter le langage moderne et travailler dans le plan muni d'un repère orthonormé usuel et ainsi travailler en coordonnées. Nous avons placé le point $\mathrm{O} = (0, 0)$ en l'origine du repère, puis le point $\mathrm{A} = (a, 0)$, le point $\mathrm{B} = (a, b)$ ainsi que les point $\mathrm{A}' = (a', 0)$ et $\mathrm{B}' = (a', b')$. Prendre des multiples du segment de grandeur a revient simplement à considérer dans le plan tous les points de coordonnées $(ma, 0)$ pour $m \in \mathbf{Z}$. De même pour les segments de grandeurs b, a', b'. Le procédé des équimultiples de la définition 6 nous amène à considérer tous les points du plan de la forme (ma, nb), avec (m, n) variant dans $\mathbf{Z} \times \mathbf{Z}$, formant ainsi une première grille rectangulaire que l'on appelle un réseau. De même on forme les points de coordonnées des multiples entiers de a' et b'. Nous noterons les réseaux obtenus $\mathrm{R}_{a,b}$ et $\mathrm{R}_{a',b'}$:

$$\mathrm{R}_{a,b} = \big\{(ma, nb) \in \mathbf{R}^2, m \in \mathbf{Z}, n \in \mathbf{Z}\big\},\ \mathrm{R}_{a',b'} = \big\{(ma', nb') \in \mathbf{R}^2, m \in \mathbf{Z}, n \in \mathbf{Z}\big\}.$$

La définition 6 énonce alors l'égalité des rapports a/b et a'/b' en fonction des positions des points des deux réseaux $\mathrm{R}_{a,b}$ et $\mathrm{R}_{a',b'}$ par rapport à la *diagonale*. La diagonale est la droite Δ du plan d'équation $y = x$. Elle représente la *proportion triviale* d'égalité $x/y = 1/1$. Les deux proportions a/b et a'/b' sont égales, d'après la proposition 6 si, pour tout choix de couple d'entiers $(m, n) \in \mathbf{Z}$, les points correspondants (ma, nb) et (ma', nb') dans les réseaux $\mathrm{R}_{a,b}$ et $\mathrm{R}_{a',b'}$ respectivement, sont en même position par rapport à la diagonale : soit tous les deux au-dessus de Δ, soit tous les deux au-dessous, soit tous les deux exactement sur la diagonale. Dans la figure 1.2, les deux points $(2a, 3b)$ et $(2a', 3b')$ ne sont pas situés du même côté de la diagonale, exprimant le fait que $a/b \neq a'/b'$.

La proportion a/b définit de manière univoque la droite $\mathrm{D}_{a/b}$ passant par O et le point (a, b). Cette proportion n'est autre que ce que l'on appelle aujourd'hui la *pente* de la droite $\mathrm{D}_{a/b}$. Réciproquement, toute droite non verticale D passant par O définit

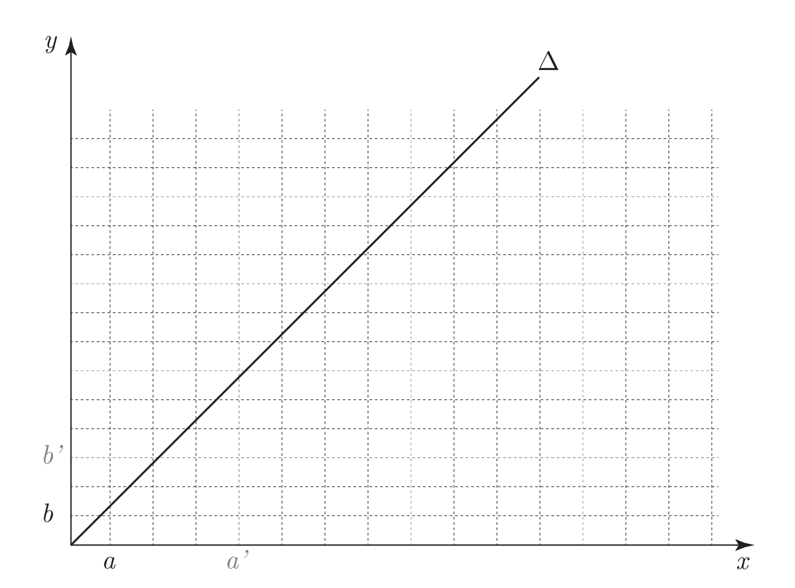


Figure 1.2 – Les réseaux $\mathrm{R}_{a,b}$, $\mathrm{R}_{a',b'}$ et la diagonale Δ

sa pente, qui n'est autre que la proportion en laquelle sont les coordonnées d'un point quelconque autre que O choisi sur icelle.

Nous allons maintenant adopter le point de vue des pentes et construire ainsi $\mathbf{R}$ au sens d'Eudoxe.

La proportion triviale $x/y = 1/1$, donnant la diagonale de pente 1, engendre également, par le procédé des équimultiples, un réseau, celui des points à coordonnées exactement entières :

$$\mathrm{R}_{1,1} = \big\{(m,n) \in \mathbf{R}^2, m \in \mathbf{Z}, n \in \mathbf{Z}\big\},$$

qui est un réseau à maille carrée de taille unité. Si l'on trace dans le plan une droite passant par l'origine $(0,0)$, deux situations peuvent se produire : soit la droite repasse par un point du réseau autre que l'origine, soit non.

Dans le premier cas il est facile de voir que la pente est alors nécessairement un nombre rationnel. Sinon la pente est irrationnelle. En guise d'exercice, vous pouvez démontrer ce qui est censé être « facile » dans les phrases qui précèdent.

Si la droite ne passe pas par l'origine, peut-on encore caractériser la rationalité de sa pente en observant sa position par rapport aux points du réseau entier ? La réponse est oui et est donnée dans la définition d'Euclide.

Donnons-nous une D non verticale passant par O, c'est-à-dire que nous supposons que la droite est de pente finie. Renversons le point de vue de la définition 6, qui observe les positions des points des réseaux rectangulaires par rapport à la diagonale, en considérant maintenant les points du réseau carré par rapport à la droite D. Cette

définition nous suggère de regarder, pour chaque entier $n \in \mathbf{Z}$, le nombre de points à coordonnées entières compris entre le point $(n, 0)$ et la droite D. Si la droite D est au-dessus de l'axe des abscisses nous compterons ce nombre positivement, et si elle est en dessous, nous le compterons négativement. Nous noterons $f_{\mathrm{D}}(n)$ ce nombre, de sorte que $f_{\mathrm{D}}(n) \in \mathbf{Z}$. Ainsi à chaque droite non verticale D nous avons associée une application $f_{\mathrm{D}} : \mathbf{Z} \longrightarrow \mathbf{Z}$ qui caractérise la position de cette droite par rapport au réseau des points à coordonnées entières (*voir* la figure 1.3).

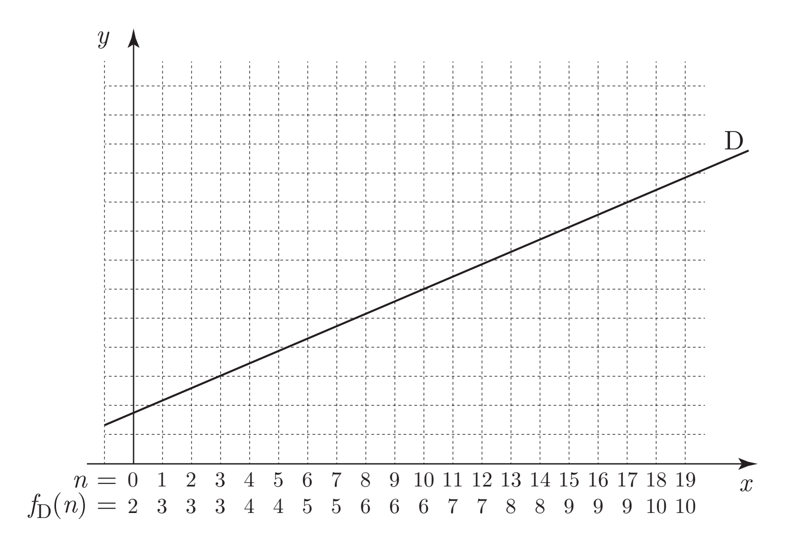


Figure 1.3 – La fonction f_{D} pour une droite de pente (semble-t-il) irrationnelle

À titre d'exercice, vous pouvez démontrer que les points de vue « position des points du réseau rectangulaire de proportion a/b par rapport à la diagonale » et « position des points du réseau carré par rapport à la droite D de pente a/b » sont équivalents.

La construction de f_{D} décrite plus haut est tout autant valide si D ne passe pas par l'origine. Il est alors aisé de démontrer (encore un exercice pour le lecteur courageux) que deux droites non verticales D et D′ sont parallèles, c'est-à-dire ont la même pente, si et seulement si les deux applications f_{D} et $f_{\mathrm{D}'}$ ont une différence *bornée* :

$$\mathrm{D} \parallel \mathrm{D}' \text{ si et seulement s'il existe } \mathrm{C} \geqslant 0 \text{ tel que } \forall n \in \mathbf{Z},\ |f_{\mathrm{D}}(n) - f_{\mathrm{D}'}(n)| \leqslant \mathrm{C}.$$

Les applications f_{D} sont par ailleurs particulières parmi toutes les applications de $\mathbf{Z}$ dans $\mathbf{Z}$. Construite à partir de la *droite* D, il est facile de voir que l'on n'aura jamais $f_{\mathrm{D}}(n) = n^2$ pour tout n par exemple. En fait f_{D} est quasiment un *morphisme* du groupe additif $\mathbf{Z}$ dans lui-même : il existe une constante $\mathrm{C} > 0$ telle que, pour tous $n, m \in \mathbf{Z}$, on ait $|f_{\mathrm{D}}(m+n) - (f_{\mathrm{D}}(m) + f_{\mathrm{D}}(n))| \leqslant \mathrm{C}$. Si C était nulle, f_{D} serait exactement un morphisme. Nous allons adopter la définition suivante :

Définition 1.11 — *Un quasi-morphisme de* $\mathbf{Z}$ *dans lui-même est une application* $f : \mathbf{Z} \longrightarrow \mathbf{Z}$ *telle qu'il existe une constante* $\mathrm{C} > 0$ *satisfaisant à*

$$\forall n, m \in \mathbf{Z},\ |f(n+m) - f(m) - f(n)| \leqslant \mathrm{C}.$$

Dit autrement, à une fonction bornée près, f *est un morphisme de* $\mathbf{Z}$ *dans lui-même.*

Toute fonction de la forme f_{D} est un quasi-morphisme et la réciproque est quasiment vraie (*sic*) au sens où si f est un quasi-morphisme, il existe une droite D telle que $f - f_{\mathrm{D}}$ soit une application bornée. L'avantage de passer des applications de la forme f_{D} aux quasi-morphismes est que ces derniers offrent un peu plus de souplesse et surtout qu'ils forment, dans leur ensemble, un groupe abélien : si f, g sont deux quasi-morphismes, alors l'application $f + g$ définie par $(f+g)(n) = f(n) + g(n)$ pour tout $n \in \mathbf{Z}$ est encore un quasi-morphisme et l'application $-f$ définie par $(-f)(n) = -f(n)$ en est un aussi. Nous noterons QM l'ensemble de tous les quasi-morphismes de $\mathbf{Z}$ dans $\mathbf{Z}$. C'est, muni de l'addition définie ci-dessus, un groupe commutatif. L'un de ses sous-groupes intéressants est celui des quasi-morphismes *bornés.* Nous noterons QM_b l'ensemble des quasi-morphismes bornés. C'est un sous-groupe de QM et $f_{\mathrm{D}} \in \mathrm{QM}_b$ si et seulement si D est horizontale, c'est-à-dire de pente nulle. Du coup on peut énoncer en nos termes savants la proposition caractérisant le parallélisme de deux droites :

$$\mathrm{D} \parallel \mathrm{D}' \text{ si et seulement si } f_{\mathrm{D}} - f'_{\mathrm{D}} \in \mathrm{QM}_b.$$

Mais nous reconnaissons-là la définition de l'égalité de f_{D} et f'_{D} modulo le sous-groupe QM_b. Ainsi devient-il clair qu'il doit y avoir correspondance bi-univoque entre les pentes des droites et les classes de congruences de quasi-morphismes modulo quasi-morphisme borné.

Définition 1.12 — *Un nombre réel est un élément du groupe abélien quotient* $\mathrm{QM}/\mathrm{QM}_b$*. On note ce quotient* $\mathbf{R}$.

Nous disposons, avec cette définition, d'une addition naturelle, celle du groupe quotient. Pour la notion d'ordre, nous allons nous appuyer sur l'intuition géométrique : une droite est de pente strictement positive si elle monte. Cela signifie en particulier que, quitte à se placer suffisamment loin à droite dans le plan, la droite va passer au-dessus de l'axe des abscisses et y rester. Mieux, comme elle monte strictement, elle va passer au-dessus de n'importe quelle droite horizontale, aussi haute celle-ci se situe-t-elle.

Définition 1.13 — *On dit qu'un quasi-morphisme* $f \in \mathrm{QM}$ *est positif si pour tout* $\mathrm{C} \in \mathbf{Z}$ *il existe* $\mathrm{N} \in \mathbf{Z}$ *tel que, pour tout* $n \geqslant \mathrm{N}$ *on ait* $f(n) > \mathrm{C}$.

On peut démontrer que la notion de positivité passe au quotient et permet d'ordonner les nombres réels.

Pour définir la multiplication des nombres réels, nous allons encore une fois faire appel à la géométrie. Soient en effet les deux droites d'équations $y = ax$ et $y = a'x$ respectivement. Leurs pentes respectives valent a et a'. Elles représentent les graphes des applications $f(x) = ax$ et $g(x) = a'x$. Si l'on compose ces deux applications on trouve $(f \circ g)(x) = aa'x$, dont le graphe est la droite de pente aa'. Voilà !

Définition 1.14 — *Si f, g sont deux quasi-morphismes, on définit leur multiplication par $f \times g = f \circ g$.*

Cette multiplication passe au quotient $\mathbf{R} = \mathrm{QM}/\mathrm{QM}_b$ et nous voilà munis du corps des réels : nous laissons la lectrice courageuse démontrer que $\mathrm{QM}/\mathrm{QM}_b$ est bien un corps, totalement ordonné, archimédien et satisfaisant à la propriété des coupures.

Remarque : l'idée de considérer des quasi-morphismes de groupes est due, je crois, à Michael Gromov, dans ses travaux sur la *cohomologie bornée.* On trouvera une exposition complète de la construction des nombres réels *à la* Eudoxe dans un article du mathématicien Norbert A'Campo disponible sur internet à l'adresse *http://arxiv.org/abs/math/0301015.*

Chapitre 2
Des aires aux primitives, et *vice versa*

Nous savons tous calculer l'aire de figures simples, triangles ou rectangles. Pour des polygones plus compliqués, nous prenons des ciseaux et essayons de nous ramener à un nombre fini de petits bouts simples (procédé utile pour les cadastres, par exemple) ; il s'agit finalement d'une méthode d'approximation, et nous allons raffiner cette idée pour les régions plus compliquées traitées par la définition 1 mentionnée dans l'introduction.

2.1. Intégrale et aire

Soit f une fonction disons, pour simplifier, positive et continue sur l'intervalle $[a, b]$. On prétend dans le théorème 1 de l'introduction calculer l'aire A du domaine R contenu sous le graphe de f.

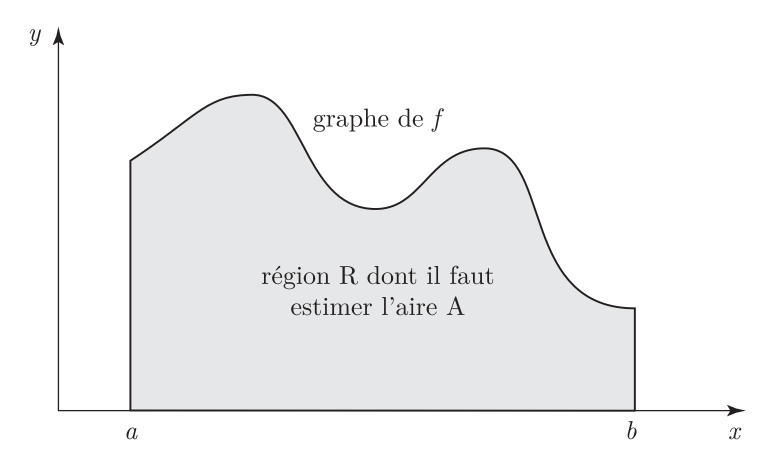


Figure 2.1 – L'intégrale comme aire d'une région plane

L'idée la plus simple, pour calculer cette aire de manière approchée, est de découper la région R en rectangles verticaux, choisis, dans un premier temps *au pif.* Soyons simplistes et un peu fainéants : contentons nous de trois rectangles, basés respectivement sur les intervalles $\mathrm{I}_1 = [a = x_0, x_1], \mathrm{I}_2 = [x_1, x_2], \mathrm{I}_3 = [x_2, x_3 = b]$. Choisissons un point dans chacun de ces intervalles : $t_1 \in \mathrm{I}_1, t_2 \in \mathrm{I}_2$ et $t_3 \in \mathrm{I}_3$, et dessinons des rectangles de base I_j, de hauteur $f(t_j)$ (*voir* la figure 2.2).

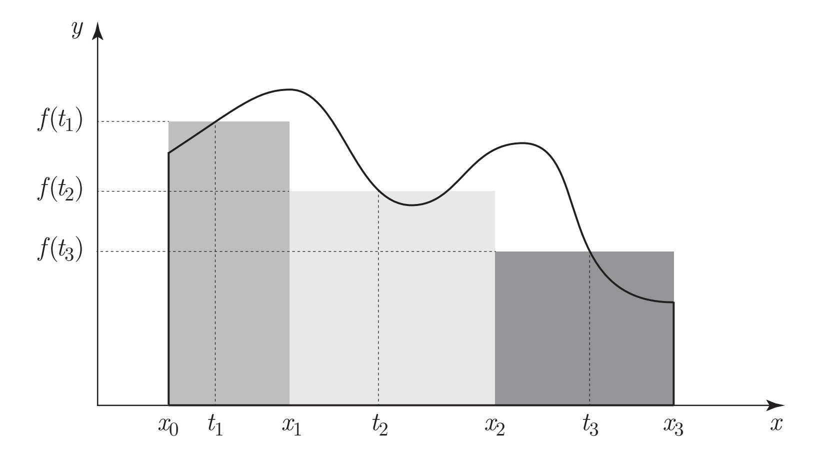


Figure 2.2 – Approximation de l'aire A par des rectangles

Nous obtenons ainsi trois rectangles d'aires respectives

$$\mathrm{A}_1 = (x_1 - x_0)f(t_1),\ \mathrm{A}_2 = (x_2 - x_1)f(t_2) \text{ et } \mathrm{A}_3 = (x_3 - x_2)f(t_3),$$

soit une approximation de l'aire de la région R par la somme

$$\mathrm{A} \approx \mathrm{A}_1 + \mathrm{A}_2 + \mathrm{A}_3 = \sum_{k=1}^{3}(x_k - x_{k-1})f(t_k) = \sum_{k=1}^{3} |\mathrm{I}_k| f(t_k),$$

où l'on a noté $|\mathrm{I}_k|$ la *longueur* de l'intervalle I_k. Nous venons de définir deux notions qui seront centrales dans toute la suite de ce cours, à savoir celle de *subdivision pointée* (la famille des intervalles $\mathrm{I}_1, \mathrm{I}_2, \mathrm{I}_3$ et le choix d'un point t_j dans chacun des I_j), et celle de *somme de Riemann* (la somme $\mathrm{A}_1 + \mathrm{A}_2 + \mathrm{A}_3$) associée à cette subdivision.

Définition 2.1 — *Soit* $\mathrm{I} = [a, b]$ *un intervalle compact. Une subdivision pointée* $\mathcal{P}$ *de* I *est la donnée d'une subdivision* $\mathrm{I}_1 = [x_0, x_1], \mathrm{I}_2 = [x_1, x_2], \ldots, \mathrm{I}_n = [x_{n-1}, x_n]$ *de* I *(avec* $a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n = b$*) et d'un « pointage » de cette partition, c'est-à-dire de points* $t_1 \in \mathrm{I}_1, \ldots, t_n \in \mathrm{I}_n$*. On la notera* $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$*, et on appellera les* t_i *des points de marquage de* $\mathcal{P}$*.*

Notons au passage que les subdivisions que nous considérerons seront toujours supposées ordonnées par l'ordre usuel de $\mathbf{R}$. Les subdivisions pointées de l'intervalle $[a, b]$

permettent naturellement, comme nous venons de le voir, d'approcher l'aire du domaine A sous le graphe d'une fonction f définie sur $[a, b]$.

Définition 2.2 — *Soient f une fonction numérique* quelconque *définie sur l'intervalle $[a, b]$, et $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$ une subdivision pointée de $[a, b]$. On appelle somme de Riemann de f associée à $\mathcal{P}$ la quantité :*

$$\mathrm{S}(f, \mathcal{P}) = \sum_{k=1}^{n} |\mathrm{I}_k| f(t_k).$$

Nous disposons maintenant d'un outil simple d'approximation des intégrales au sens de la définition 1, et il serait agréable que cette méthode d'approximation permette de faire le lien entre intégrale et primitive comme cela est énoncé dans le théorème 1. Voyons de plus près ce qu'il en est et essayons d'estimer les sommes de Riemann d'une *fonction dérivée.*

2.2. Intégrale et primitive

Considérons une fonction f, dérivable sur $[a, b]$ (en a et b, il s'agit de dérivées à droite et à gauche respectivement) et essayons d'approcher l'intégrale de f' par des sommes de Riemann, comme la section précédente nous le suggère. Donnons-nous une subdivision pointée $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$ de $[a, b]$, et formons la somme de Riemann de f' associée à $\mathcal{P}$:

$$\mathrm{S}(f', \mathcal{P}) = \sum_{k=1}^{n} (x_{k-1} - x_k) f'(t_k),$$

où, bien entendu, on a noté $\mathrm{I}_k = [x_{k-1}, x_k]$. La quantité $(x_{k-1} - x_k) f'(t_k)$ n'est pas sans rappeler la formule des accroissements finis qui stipule que, sous les hypothèses faites ici, il existe un point $c_k \in [x_{k-1}, x_k]$ tel que

$$f(x_k) - f(x_{k-1}) = (x_k - x_{k-1}) f'(c_k).$$

En termes plus géométriques, la tangente au graphe de f au point $(c_k, f(c_k))$ est parallèle à la *corde* entre les points $(a, f(a))$ et $(b, f(b))$ (*voir* la figure 2.3).

Si, par le plus grand des hasards et grâce à une chance incroyable, nous avions $t_k = c_k$ pour tous les k, la somme de Riemann de f' associée à $\mathcal{P}$ deviendrait une *somme télescopique* se simplifiant par miracle :

$$\begin{aligned} \mathrm{S}(f', \mathcal{P}) &= (x_0 - x_1) f'(c_1) + (x_1 - x_2) f'(c_2) + \cdots + (x_n - x_{n-1}) f'(c_n) \\ &= (f(x_1) - f(x_0)) + (f(x_2) - f(x_1)) + \cdots + (f(x_n) - f(x_{n-1})), \end{aligned}$$

soit :

$$\mathrm{S}(f', \mathcal{P}) = f(x_n) - f(x_0) = f(b) - f(a).$$

Nous aurions ainsi trouvé *exactement* la formule du théorème 1 ! Bien sûr, comme nous parlons d'approximation, il est insensé de croire que la situation $c_k = t_k$ puisse réellement avoir lieu, sauf dans le cas où f est affine.

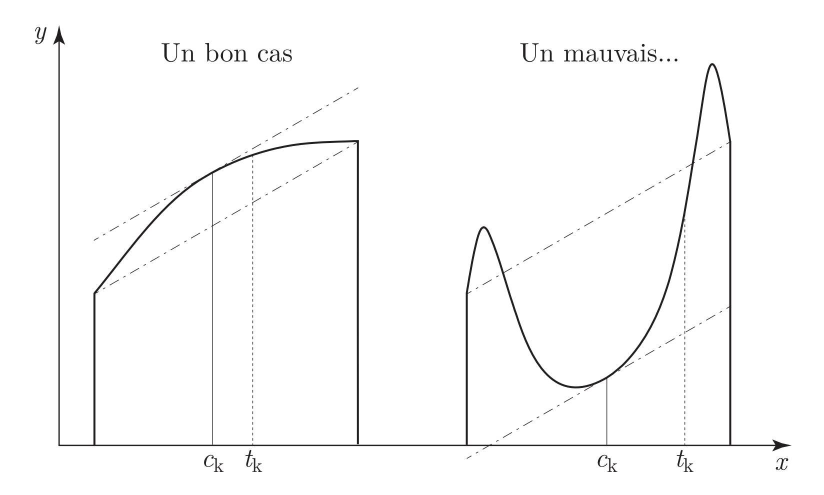


Figure 2.3 – Deux cas de la formule des accroissements finis

Dans tous les autres cas, nous ne pouvons espérer qu'une seule chose : qu'en prenant les intervalles I_k suffisamment petits, les points c_k et t_k soient tellement proches que l'on ait alors $f'(c_k) \approx f'(t_k)$. Cela donne une première idée pour définir l'intégrale de f' : ce serait la limite, si elle existe, des sommes de Riemann $S(f', \mathcal{P})$, quand on prend des subdivisions de pas de plus en plus petit. On pourrait par exemple prendre pour sommets des intervalles les nombres du type $a + j(b-a)2^{-l}, j = 0, \ldots, 2^l$, c'est à dire découper $[a, b]$ en 2^l intervalles de longueur égale, et ensuite faire tendre l vers l'infini. Mais ce n'est, *a priori*, pas suffisant : cela est illustré par la figure 2.3. Dans le premier cas, le « bon », prendre t_k différent de c_k donnera une bonne approximation de

$$(x_k - x_{k-1})f'(c_k) = f(x_k) - f(x_{k-1})$$

par

$$(x_k - x_{k-1})f'(t_k),$$

car f' (c.-à-d. la pente du graphe) varie peu dans l'intervalle. Mais, dans le second, un choix malencontreux de t_k (par exemple en un point où la dérivée de f prend une valeur beaucoup plus grande que l'accroissement entre x_{k-1} et x_k comme c'est le cas sur la figure ci-dessus) donnera une approximation déplorable. Il existe des fonctions f dérivables, dont la dérivée varie au voisinage de certains points si fortement que l'usage d'intervalles de longueur constante ne permet pas de faire converger les sommes de Riemann vers une limite. Il faut assouplir le processus de passage à la limite et s'autoriser à *adapter* la longueur des intervalles des subdivisions en fonction des variations locales de f'. Ces considérations nous mènent tout naturellement aux deux définitions suivantes, celle de *jauge* et celle de subdivision *δ-fine* :

Définition 2.3 — *Un pas, ou* jauge, *est une fonction δ définie sur $[a,b]$ et à valeurs dans $]0,+\infty[$. Une subdivision pointée $\mathcal{P}=\{(\mathrm{I}_1,t_1),\ldots,(\mathrm{I}_n,t_n)\}$ sera dite adaptée au pas, ou encore δ-fine, si, pour tout $1\leqslant k\leqslant n$ on a*

$$\mathrm{I}_k\subseteq\left[t_k-\frac{\delta(t_k)}{2},t_k+\frac{\delta(t_k)}{2}\right].$$

On remarque en particulier que $|\mathrm{I}_k|\leqslant\delta(t_k)$.

Un exemple simple de jauge est donné, pour tout *nombre* strictement positif δ, par la jauge constante valant partout δ et que nous noterons encore δ.

Le but du jeu est maintenant de trouver une jauge δ, qui nous assure que l'on a une bonne approximation de $f(b)-f(a)$ par $\mathrm{S}(f',\mathcal{P})$ dès que la subdivision $\mathcal{P}$ est δ-fine. Commençons par nous donner un $\varepsilon>0$, éventuellement petit, par exemple $\varepsilon=1/1000$. C'est l'ordre de précision que nous exigeons sur l'approximation de l'intégrale par les sommes de Riemann : peut-on trouver δ de manière à ce que, si $\mathcal{P}$ est δ-fine, alors $\mathrm{S}(f',\mathcal{P})$ est égal à $f(b)-f(a)$ à moins de ε près ? La discussion qui précède nous indique qu'une bonne jauge devrait permettre de dire que l'on est dans le « bon cas » de la formule des accroissements finis sur chaque I_k d'une subdivision δ-fine.

Une jauge qui satisferait à la condition suivante (notons cette condition $(\star)$) serait *au poil* : pour tout $x\in[a,b]$, et tout $u\neq v$ satisfaisant

$$x-\delta(x)/2\leqslant u\leqslant x\leqslant v\leqslant x+\delta(x)/2,$$

on a

$$(\star)\qquad |f(v)-f(u)-f'(x)(v-u)|<\frac{\varepsilon}{b-a}(v-u).$$

En effet, si $\mathcal{P}$ était une partition δ-fine pour une telle jauge δ, en remarquant que

$$f(b)-f(a)=\sum_{k=1}^{n}\left(f(x_k)-f(x_{k-1})\right),$$

un petit coup de télescope et l'inégalité triangulaire donneraient :

$$\begin{aligned}|f(b)-f(a)-\mathrm{S}(f',\mathcal{P})|&=\left|\sum_{k=1}^{n}\left(f(x_k)-f(x_{k-1})\right)-\sum_{k=1}^{n}(x_k-x_{k-1})f'(t_k)\right|\\&\leqslant\sum_{k=1}^{n}\big|f(x_k)-f(x_{k-1})-f'(t_k)(x_k-x_{k-1})\big|,\end{aligned}$$

soit, du fait que δ satisfait l'hypothèse $(\star)$,

$$|f(b)-f(a)-\mathrm{S}(f',\mathcal{P})|\leqslant\varepsilon.$$

Nous voyons que nous approchons du but : d'une part définir une intégrale à partir d'approximations d'aires, en utilisant des sommes de Riemann et des subdivisions à

pas adapté, et d'autre part démontrer que cette intégrale vérifie, sous des hypothèses très générales, le théorème 1.

A-t-on espoir de trouver une jauge δ satisfaisant l'hypothèse $(\star)$? Si $x \in [a, b]$, f est dérivable en x. Ainsi

$$\lim_{t \to x} \frac{f(t) - f(x)}{t - x} = f'(x).$$

En réécrivant cette identité avec des *epsilon* et des *delta,* on obtient l'existence d'un $\delta(x) > 0$ tel que, pour tout t satisfaisant $|t - x| < \delta(x)$, on ait

$$\left| \frac{f(t) - f(x)}{t - x} - f'(x) \right| < \varepsilon' = \frac{\varepsilon}{b - a},$$

soit

$$|f(t) - f(x) - f'(x)(t - x)| < \varepsilon' |t - x|.$$

Soient alors u, v comme dans la condition $(\star)$. Le théorème[1] $a = a + b - b$ et l'inégalité triangulaire nous permettent d'écrire :

$$\begin{aligned} & |f(v) - f(u) - f'(x)(v - u)| \\ = \ & |[f(v) - f(x) - f'(x)(v - x)] - [f(u) - f(x) - f'(x)(u - x)]| \\ \leqslant \ & |f(v) - f(x) - f'(x)(v - x)| + |f(u) - f(x) - f'(x)(u - x)| \\ < \ & \varepsilon'(v - x) + \varepsilon'(x - u) = \frac{\varepsilon}{b - a}(v - u). \end{aligned}$$

Voilà, nous avons trouvé une jauge δ satisfaisant $(\star)$, et nous sommes maintenant en mesure de passer à la définition générale de fonction intégrable et d'intégrale, *via* les sommes de Riemann et les subdivisions δ-fines. Rendez-vous au chapitre suivant !

2.3. Appendice : calcul numérique d'une primitive par la méthode d'Euler

Nous n'aborderons que très peu les méthodes de calcul numérique d'intégrales dans ce cours, et renvoyons le lecteur aux traités classiques sur le sujet, comme le livre de Jean-Pierre Demailly *Analyse numérique et équations différentielles* paru chez EDP Sciences dans la collection Grenoble Sciences. Nous présentons cependant ici l'algorithme le plus naïf qui soit pour calculer une approximation d'une primitive : la *méthode d'Euler.* Celle-ci est en effet très proche de la théorie des jauges de Kurzweil et Henstock, dont elle est en quelque sorte l'avatar numérique. On pourrait très bien, pour introduire l'intégrale, partir de la méthode d'Euler et montrer comment, en la poussant dans ses derniers retranchements théoriques, elle mène à l'intégrale que nous avons présentée ici. Notons que la méthode d'Euler est en fait utilisée et présentée,

1. Le meilleur ami du mathématicien !

en général, pour le calcul numérique approché de solutions d'équations différentielles, et non simplement de primitives. Le cas que nous traitons ici est donc très particulier mais il nous semble instructif.

Le problème est le suivant : étant donné une fonction f définie sur $[a,b]$, peut-on déterminer une approximation d'une solution au problème suivant :

$$(\dagger) \qquad \begin{cases} \mathrm{F}'(t) = f(t), \ t \in [a,b] \\ \mathrm{F}(a) = y_0, \end{cases}$$

où y_0 est la condition initiale ou condition au bord du problème (†). Comme il s'agit de calcul numérique, nous ne pouvons calculer une solution approchée que sur un nombre fini de points. Ces points $x_0 = a < x_1 < \cdots < x_\mathrm{N} = b$ forment une subdivision de $[a,b]$. On pose $h_i = x_{i+1} - x_i$, $\mathrm{H} = \max\{h_i, i = 0,\ldots,\mathrm{N}-1\}$ et $h = \min\{h_i\}$. Une bonne méthode devrait assurer qu'à mesure que H tend vers 0, alors toute solution approximative calculée sur les points x_i s'approche des valeurs d'une vraie solution. Enfin, pour des raisons pratiques de précision et de temps de calcul, nous devons écrire un algorithme où h est minoré par une constante, petite, dépendant de la machine sur laquelle on travaille. L'idée va être de construire pas à pas les points x_i, en les prenant plus proches quand l'erreur risque d'augmenter, et plus éloignés quand on est sûr que l'erreur diminue. Mettons tout cela en ordre.

Nous supposerons f dérivable, même si elle peut, par endroit, présenter de brusques sautes d'humeur. Soit Ψ une solution exacte de (†), et supposons-nous donnée la subdivision $\{x_i\}$. Un développement de Taylor de Ψ au second ordre donne

$$\Psi(x_{i+1}) = \Psi(x_i + h_i) = \Psi(x_i) + h_i f(x_i) + \frac{1}{2} h_i^2 f'(x_i) + o(h_i^2).$$

La méthode d'Euler consiste à prendre pour solution approchée la suite des valeurs donnée par

$$\begin{cases} \mathrm{F}(a) = y_0 \\ \mathrm{F}(x_{i+1}) = \mathrm{F}(x_i) + h_i f(x_i). \end{cases}$$

Dans la pratique, on ne peut pas prendre exactement $\mathrm{F}(a) = y_0$, mais nous verrons plus loin que ce problème n'en est pas vraiment un.

On appelle *erreur de consistance* de la méthode à l'étape i, le réel

$$\mathrm{e}_i = \Psi(x_{i+1}) - \mathrm{F}(x_{i+1}),$$

si l'on suppose que $\mathrm{F}(x_i) = \Psi(x_i)$. En reportant dans le développement de Taylor ci-dessus, on trouve

$$\mathrm{e}_i = \frac{h_i^2}{2} f'(x_i) + o(h_i^2).$$

Ainsi, plus f varie violemment au voisinage de x_i (c.-à-d. $|f'(x_i)|$ grand) plus il va falloir prendre h_i petit et donc augmenter le nombre de points d'évaluation, et ainsi le temps de calcul. En revenant à la définition de e_i (en se rappelant qu'on suppose $\mathrm{F}(x_i) = \Psi(x_i)$), on a

$$\mathrm{e}_i = \Psi(x_{i+1}) - \Psi(x_i) - h_i f(x_i)$$

et le théorème des accroissements finis nous assure de l'existence d'un $c_i \in [x_i, x_{i+1}]$ tel que

$$\mathrm{e}_i = h_i f(c_i) - h_i f(x_i) = h_i[f(c_i) - f(x_i)].$$

Comme f est continue sur l'intervalle compact $[a, b]$, elle y est uniformément continue, et donc, si $\varepsilon > 0$ est fixé, il existe $\delta > 0$ tel que, si $\mathrm{H} < \delta$, alors

$$\sum_{i=0}^{\mathrm{N}-1} |\mathrm{e}_i| \leqslant \sum_{i=0}^{\mathrm{N}-1} h_i |f(c_i) - f(x_i)| \leqslant \varepsilon(b-a).$$

On a montré la proposition suivante :

Proposition 2.1 — *La méthode d'Euler est* consistante*, en d'autres termes, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que, si $\mathrm{H} < \delta$, alors*

$$\sum |\mathrm{e}_i| \leqslant \varepsilon.$$

Dit autrement : si la méthode converge, elle donne une approximation d'une solution exacte. Reste maintenant à savoir si des erreurs d'arrondi successives ou une erreur dans la condition initiale influent sur le résultat trouvé : c'est le problème de la *stabilité* de la méthode.

En effet, dans la pratique, on fait, à chaque étape (et ce dès la condition initiale) des erreurs d'arrondi. Supposons-nous donnée une suite d'erreurs ε_i, et donnons-nous une solution, perturbée par ces erreurs, de l'algorithme d'Euler :

$$\begin{cases} \tilde{\mathrm{F}}(x_0) = \tilde{y}_0 = y_0 + \varepsilon_0 \\ \tilde{\mathrm{F}}(x_{i+1}) = \tilde{\mathrm{F}}(x_i) + h_i f(x_i) + \varepsilon_{i+1}. \end{cases}$$

On aimerait savoir si $\tilde{\mathrm{F}}$ diffère beaucoup de F et donc évaluer

$$\tilde{\mathrm{E}} = \max_i |\tilde{\mathrm{F}}(x_i) - \mathrm{F}(x_i)|.$$

Mais $|\tilde{\mathrm{F}}(x_{i+1}) - \mathrm{F}(x_{i+1})| \leqslant |\tilde{\mathrm{F}}(x_i) - \mathrm{F}(x_i)| + |\varepsilon_{i+1}|$, et donc

$$\tilde{\mathrm{E}} \leqslant |\tilde{y}_0 - y_0| + \sum_{i=1}^{\mathrm{N}} |\varepsilon_i|,$$

et l'on a ainsi démontré la

Proposition 2.2 — *La méthode d'Euler est* stable*, de constante de stabilité $\mathrm{S} = 1$: pour toute solution perturbée $\tilde{\mathrm{F}}$ comme ci-dessus, on a*

$$\max_i |\tilde{\mathrm{F}}(x_i) - \mathrm{F}(x_i)| \leqslant \mathrm{S}\Big(|\tilde{y}_0 - y_0| + \sum_{i=1}^{\mathrm{N}} |\varepsilon_i|\Big).$$

On peut maintenant conclure cette étude générale de la méthode d'Euler, en montrant qu'elle est *convergente* :

Théorème 2.3 — *Pour toute solution exacte* Ψ *de* (†) *et toute solution approchée donnée par la méthode d'Euler :*

$$\begin{cases} \mathrm{F}(a) = \tilde{y}_0 \\ \mathrm{F}(x_{i+1}) = \mathrm{F}(x_i) + h_i f(x_i), \end{cases}$$

l'erreur globale E *vérifie*

$$\mathrm{E} = \max_i |\mathrm{F}(x_i) - \Psi(x_i)| \longrightarrow 0$$

quand $|\tilde{y}_0 - y_0|$ *et* H *tendent vers* 0. *On dit que la méthode est* convergente.

Cela découle immédiatement du fait que la méthode est consistante et stable.

Finissons cette section par une possible mise en pratique sur machine. Comme nous l'avons vu lors du calcul de l'erreur de consistance e_i, il faudrait adapter à chaque étape le pas h_i de notre subdivision pour contrer les possibles variations brusques de f. Les conditions de mise en pratique (précision de la machine, temps de calcul...) nous fixent un intervalle de contrôle $[h_{\min}, h_{\max}]$ dans lequel le pas devra se trouver. Fixons-nous par ailleurs $\varepsilon > 0$, qui va être la précision que l'on exige sur la solution approchée. On doit commencer par pouvoir prendre $\tilde{y}_0$ à moins de $\varepsilon/2$ de y_0. Ensuite, on choisit h_{i+1} de telle sorte que

$$\mathrm{e}_i^* = \frac{1}{2}(f(x_{i+1}) - f(x_i)) \approx \mathrm{e}_i/h_i$$

vérifie

$$|\mathrm{e}_i^*| \leqslant \frac{\varepsilon}{2(b-a)}.$$

On obtient ainsi une erreur globale majorée, grossièrement, par ε. Remarquons que comme on calcule les $f(x_i)$, le calcul de e_i^* ne perturbe pas beaucoup le temps de moulinage du programme.

Si l'erreur e_i^* est petite, on augmente le pas, si elle est grande, on le diminue, et si elle explose en imposant que h_{i+1} sorte de l'intervalle de contrôle, on arrête le programme. On pourrait par exemple faire comme suit :

– si $\varepsilon/(3(b-a)) \leqslant |\mathrm{e}_i^*| \leqslant \varepsilon/(2(b-a))$, on pose $h_{i+1} = h_i$;

– si $|\mathrm{e}_i^*| < \varepsilon/(3(b-a))$, on pose $h_{i+1} = \max\{5h_i/4, h_{\max}\}$;

– si $|\mathrm{e}_i^*| > \varepsilon/(b-a)$, on pose $h_{i+1} = 4h_i/5$,

et si dans la procédure on obtient $h_{i+1} < h_{\min}$, on arrête tout.

This page intentionally left blank

Chapitre 3
Fonctions intégrables, intégrale

Nous avons vu, dans le chapitre précédent, les notions de subdivision pointée, de somme de Riemann, ainsi que de subdivision δ-fine pour une jauge δ sur $[a, b]$. Nous sommes maintenant en mesure de définir la notion d'intégrale et d'intégrabilité :

Définition 2 — *Soit f une fonction numérique définie sur $[a, b]$. La fonction f est dite intégrable sur $[a, b]$ s'il existe un nombre réel S tel que, pour tout $\varepsilon > 0$, il existe une jauge δ_ε sur $[a, b]$ telle que, pour toute subdivision δ_ε-fine $\mathcal{P}$, on ait*

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}| < \varepsilon.$$

On notera $\mathcal{I}([a, b])$ l'ensemble des fonctions intégrables sur $[a, b]$. Le nombre S ci-dessus est appelé intégrale de la fonction f sur $[a, b]$ et est noté

$$\int_{[a,b]} f(x)\,\mathrm{d}x \quad \textit{ou encore} \quad \int_a^b f(x)\,\mathrm{d}x, \quad \textit{voire} \quad \int_{[a,b]} f \quad \textit{ou} \quad \int_a^b f.$$

La stratégie pour démontrer qu'une fonction donnée est intégrable est alors claire : on se fixe un réel $\varepsilon > 0$ et l'on cherche à construire la jauge δ_ε. Nous allons voir dans un instant, sur des exemples, que cela marche très bien pour un grand nombre de fonctions, comme les polynômes par exemple. Mais il nous faut tout d'abord régler un « détail » : si jamais il existait une jauge δ n'admettant pas de subdivision δ-fine, alors toute fonction serait intégrable et n'importe quelle nombre serait son intégrale !

3.1. Le lemme de Cousin

Cette chausse-trape est heureusement évitée grâce au lemme suivant, dit lemme de Cousin :

Lemme 3.1 — *Pour toute jauge $\delta > 0$ sur $[a, b]$, il existe une subdivision δ-fine.*

DÉMONSTRATION — Nous allons faire un raisonnement par l'absurde et supposer que l'intervalle $[a,b]$ ne possède pas de subdivision δ-fine.

Posons $I_0 = [a,b]$ et coupons l'intervalle $[a,b]$ en deux parties égales $J_1 \cup J_1'$, où

$$J_1 = \left[a, \frac{a+b}{2}\right] \text{ et } J_1' = \left[\frac{a+b}{2}, b\right].$$

Par hypothèse, l'un au moins des deux intervalles J_1 ou J_1' au moins ne possède pas de subdivision δ-fine. En effet, si $\mathcal{P}_1$ (resp. $\mathcal{P}_1'$) est une subdivision δ-fine de J_1 (resp. de J_1'), alors $\mathcal{P}_1 \cup \mathcal{P}_1'$ serait une subdivision δ-fine de $[a,b]$, qui n'en admet pas par hypothèse. Nous définirons donc I_1 comme l'un des deux intervalles J_1, J_1' qui n'admet pas de subdivision δ-fine. Ce procédé peut s'itérer : on coupe I_1 en deux parties égales et l'on choisit pour $I_2 \subset I_1$ l'une de ces parties qui n'admet pas de subdivision δ-fine. Bien sûr, la longueur de I_1 est moitié de celle de I_0 et celle de I_2 moitié de celle de I_1. Continuant ainsi, nous obtenons une suite $(I_n)_{n\geqslant 0}$ d'intervalles fermés qui satisfont les propriétés suivantes :

- pour tout n, l'intervalle I_n ne possède pas de subdivision δ-fine ;
- les intervalles forment une suite décroissante d'intervalles emboîtés, c'est-à-dire que $I_{n+1} \subset I_n$;
- la longueur des I_n tend vers zéro quand n tend vers l'infini, puisque $|I_n| \leqslant 2^{-n}|I_0|$.

D'après le théorème des segments emboîtés, il existe un unique réel x qui soit élément de tous les I_n à la fois. En d'autres termes,

$$\bigcap_{n\in\mathbf{N}} I_n = \{x\}.$$

Comme δ est une jauge, on a $\delta(x) > 0$, et donc il existe, d'après la propriété d'Archimède, un entier $n \geqslant 0$ tel que $2^{-n}|I_0| < \delta(x)/2$. Mais alors la subdivision triviale $\mathcal{P} = \{(I_n, x)\}$ est une subdivision δ-fine de l'intervalle I_n, ce qui est impossible puisqu'il n'en possède pas. Cela est la contradiction recherchée, et le lemme de Cousin est démontré. □

Ouf ! Nous venons de voir, pour la première fois, le besoin d'un argument de topologie de la droite réelle. Ce n'est pas la dernière ! Le lemme de Cousin, malgré une preuve simple, est d'une grande puissance. Voici un premier exemple d'application, nous en donnons quelques autres en appendice à ce chapitre.

Un exemple d'application du lemme de Cousin, le théorème de Heine-Borel

Soit F un fermé borné de $\mathbf{R}$; on peut supposer que F est inclus dans un intervalle borné $[a,b]$. Nous supposons que pour tout $x \in F$, on s'est donné un intervalle ouvert $J(x)$ qui contienne x. En d'autres termes, on s'est donné un *recouvrement*

de F par des *ouverts*. L'intervalle $\mathrm{J}(x)$ étant ouvert et non vide il est de longueur non nulle. Définissons sur $[a,b]$ une jauge de la manière suivante :

– si $x \in \mathrm{F}$, on prend $\delta(x)$ suffisamment petit pour que

$$[x - \delta(x)/2, x + \delta(x)/2] \subset \mathrm{J}(x)\,;$$

– si $x \in [a,b] - \mathrm{F}$, comme le complémentaire de F est ouvert, on peut trouver $\delta(x) > 0$ tel que

$$[x - \delta(x)/2, x + \delta(x)/2] \cap \mathrm{F} = \emptyset.$$

D'après le lemme de Cousin, il existe une subdivision δ-fine $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$ de $[a,b]$. Par construction de δ, cette subdivision jouit des deux propriétés suivantes :

– si $t_k \in \mathrm{F}$, alors $\mathrm{I}_k \subset \mathrm{J}(t_k)$;

– si $t_k \notin \mathrm{F}$, alors $\mathrm{I}_k \cap \mathrm{F} = \emptyset$.

Ainsi, F est inclus dans l'union des intervalles I_k tels que $t_k \in \mathrm{F}$, et l'on en déduit que F est inclus dans l'union des $\mathrm{J}(t_k)$ correspondants. Nous avons donc démontré que du recouvrement de F par les ouverts $\mathrm{J}(x), x \in \mathrm{F}$, on peut extraire un sous-recouvrement *fini* : c'est le théorème de Heine-Borel.

Avant de poursuivre, commençons par donner la première propriété de l'intégrale d'une fonction, à savoir son unicité : il serait en effet embêtant que l'intégrale d'une même fonction puisse avoir plusieurs valeurs possibles.

Proposition 3.2 — *Si f est intégrable sur $[a,b]$, alors son intégrale* S *est unique.*

Démonstration — C'est une conséquence immédiate de la séparation de $\mathbf{R}$, qui dit, entre autres, que si une suite converge vers l et vers l' alors $l = l'$. Supposons donc que f ait deux intégrales, S et S′. Soit $\varepsilon > 0$. Par définition de S, il existe une jauge δ telle que, pour toute subdivision δ-fine $\mathcal{P}$ on ait

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}| < \varepsilon.$$

Par ailleurs, il existe une autre jauge δ' telle que, pour toute subdivision δ'-fine $\mathcal{P}'$, on ait

$$|\mathrm{S}(f, \mathcal{P}') - \mathrm{S}'| < \varepsilon.$$

Ce qui serait bien maintenant, c'est de trouver une jauge δ'' telle que toute subdivision $\mathcal{P}''$ qui soit δ''-fine soit aussi fine pour δ et δ'. En prenant alors une telle subdivision $\mathcal{P}''$, on aurait, en additionnant les deux inégalités ci-dessus,

$$|\mathrm{S} - \mathrm{S}'| \leqslant |\mathrm{S} - \mathrm{S}(f, \mathcal{P}'')| + |\mathrm{S}(f, \mathcal{P}'') - \mathrm{S}'| < 2\varepsilon,$$

et, comme cette inégalité est vraie pour tout ε, la proposition serait démontrée : $\mathrm{S} = \mathrm{S}'$. En fait, une jauge telle que δ'' existe toujours, c'est une conséquence de l'utile lemme suivant.

Lemme 3.3 — *Soient* $\delta_1, \dots, \delta_l$ *un nombre fini de jauges sur* $[a, b]$*. Alors* $\delta = \min\{\delta_1, \dots, \delta_n\}$ *est une jauge. De plus, toute subdivision pointée* δ*-fine est aussi* δ_i*-fine pour* $1 \leqslant i \leqslant l$*.*

La démonstration est laissée en exercice, en guise d'entraînement au maniement des jauges. □

3.2. Le théorème fondamental du calcul différentiel et intégral

Revenons brièvement, dans cette section, sur notre introduction. Nous allons démontrer, comme annoncé, que l'intégrale que nous avons introduite ici répond complètement au problème du calcul inverse des dérivées. Nous ne démontrerons pas ici l'énoncé le plus général dans cette direction, même si celui-ci est accessible assez facilement. Nous nous contenterons de prouver le théorème suivant, qui est déjà une amélioration sensible de l'énoncé analogue dans les théories de Riemann et de Lebesgue.

Théorème fondamental du calcul différentiel et intégral — *Soit* $\mathrm{F} : [a, b] \longrightarrow \mathbf{R}$ *une fonction continue sur* $[a, b]$*, dérivable sur* $]a, b[$ *et admettant une dérivée à droite en* a *(resp. à gauche en* b*). Notons* f *la dérivée de* F*. Alors* f *est intégrable et l'on a*

$$\int_a^b f = \mathrm{F}(b) - \mathrm{F}(a).$$

Démonstration — La dérivabilité de F, définie par des limites ponctuelles, nous fournit une jauge naturelle. Fixons $\varepsilon > 0$. Nous avons vu au chapitre précédent que pour tout $x \in [a, b]$, il existe un $\delta(x) > 0$ tel que, pour tous $y, z \in [a, b]$ vérifiant $z - y \leqslant \delta(x)$ et $y \leqslant x \leqslant z$, on ait

$$(\dagger) \qquad |\mathrm{F}(z) - \mathrm{F}(y) - (z - y)f(x)| \leqslant \varepsilon(z - y).$$

Si $\mathcal{P}$ est une subdivision pointée δ-fine, on a par définition

$$|\mathrm{S}(f, \mathcal{P}) - (\mathrm{F}(b) - \mathrm{F}(a))| = \left| \sum_{k=1}^{n} \big(\mathrm{F}(x_k) - \mathrm{F}(x_{k-1}) - (x_k - x_{k-1})f(t_k)\big) \right|,$$

et, en sommant les inégalités (†) pour chaque terme du second membre, on trouve la majoration

$$|\mathrm{S}(f, \mathcal{P}) - (\mathrm{F}(b) - \mathrm{F}(a))| \leqslant \varepsilon \sum_{k=1}^{n} (x_k - x_{k-1}) = \varepsilon(b - a),$$

ce qui achève la démonstration du théorème. □

Nous reviendrons longuement sur les applications de ce théorème dans le chapitre consacré au calcul des primitives.

3.3. Premiers exemples de fonctions intégrables, intégrales de Riemann et de Lebesgue

Nous allons ici donner les premiers exemples de fonctions intégrables et introduire les notions de fonction intégrable au sens de Riemann et au sens de Lebesgue. La lectrice avisée peut tout à fait sauter cette section pour y revenir plus tard et passer directement à la suivante.

3.3.1. Fonctions constantes

Le lecteur avisé se convaincra aisément que la fonction nulle partout est intégrable. Fort heureusement ce n'est pas le seul exemple simple, comme le montre celui des fonctions constantes.

Proposition 3.4 — *Soit $c \in \mathbf{R}$ et notons encore c la fonction constante égale à c sur $[a, b]$. Alors c est intégrable et*

$$\int_a^b c \,\mathrm{d}x = (b-a)c.$$

C'est une évidence une fois que l'on a remarqué que les sommes de Riemann ne dépendent pas de la subdivision choisie, et valent toutes $(b-a)c$ (sommes télescopiques).

3.3.2. Fonctions polynomiales

Essayons maintenant d'intégrer un monôme de la forme $f(x) = x^n$, pour un entier $n \geqslant 0$. Bien sûr, nous connaissons le résultat : l'intégrale de x^n entre a et b vaut $\mathrm{F}(b) - \mathrm{F}(a)$, où $\mathrm{F}(x) = (n+1)^{-1}x^{n+1}$ est une primitive de f. Nous revenons donc aux idées du chapitre précédent : si $\mathcal{P} = \{([x_{k-1}, x_k], t_k), 1 \leqslant k \leqslant n\}$ est une subdivision pointée, alors la formule des accroissements finis nous garantit, pour tout k, l'existence d'un nombre $u_k \in [x_{k-1}, x_k]$ tel que

$$\mathrm{F}(x_k) - \mathrm{F}(x_{k-1}) = f(u_k)(x_k - x_{k-1}),$$

soit, en donnant à f et F leurs vraies significations,

$$\frac{1}{n+1}\left(x_k^{n+1} - x_{k-1}^{n+1}\right) = u_k^n(x_k - x_{k-1}).$$

Additionnons ces identités, et nous obtenons une somme télescopique :

$$\frac{1}{n+1}(b^{n+1} - a^{n+1}) = \sum_{k=1}^{n} u_k^n(x_k - x_{k-1}),$$

soit, en faisant la différence avec la somme de Riemann associée à $\mathcal{P}$:

$$\frac{1}{n+1}(b^{n+1}-a^{n+1})-\mathrm{S}(f,\mathcal{P})=\sum_{k=1}^{n}(u_k^n-t_k^n)(x_k-x_{k-1}).$$

Fixons $\varepsilon>0$. Il existe un nombre $\delta>0$ tel que, pour tous $x,y\in[a,b]$ vérifiant $|x-y|<\delta$, alors $|x^n-y^n|<\varepsilon$. Ceci exprime *l'uniforme continuité* de la fonction continue f sur l'intervalle compact $[a,b]$. Dans le cas de $f(x)=x^n$, nous pouvons expliciter δ en fonction de ε, en prenant par exemple,

$$\delta=\frac{\varepsilon}{\max\{n|b|^{n-1},n|a|^{n-1}\}}.$$

Si l'on note encore δ la jauge constante partout égale à δ, on peut majorer la différence ci-dessus par

$$\begin{aligned}\left|\frac{1}{n+1}(b^{n+1}-a^{n+1})-\mathrm{S}(f,\mathcal{P})\right| &\leqslant \sum_{k=1}^{n}|u_k^n-t_k^n|(x_k-x_{k-1})\\ &\leqslant \sum_{k=1}^{n}\varepsilon(x_k-x_{k-1})=\varepsilon(b-a),\end{aligned}$$

car $u_k,t_k\in\mathrm{I}_k$ et donc $|u_k^n-t_k^n|<\varepsilon$ par construction de δ. Nous pouvons conclure cette discussion par la proposition suivante :

Proposition 3.5 — *Soient $[a,b]$ un intervalle compact de $\mathbf{R}$, et $n\in\mathbf{N}$. La fonction monomiale $f(x)=x^n$ est intégrable sur $[a,b]$ et*

$$\int_a^b x^n\,\mathrm{d}x=\frac{1}{n+1}(b^{n+1}-a^{n+1}),$$

ce qui est le résultat attendu.

Nous verrons dans l'appendice du chapitre suivant, lorsque nous traiterons de la méthode de Simpson, que l'intégration des polynômes fournit des algorithmes numériques rapides et précis d'approximation des intégrales.

3.3.3. Fonctions intégrables au sens de Riemann

Dans l'exemple ci-dessus, nous avons vu qu'il suffisait, pour démontrer l'intégrabilité de x^n, de considérer des jauges constantes. Nous venons de retrouver, comme cas particulier des fonctions intégrables, les fonctions *intégrables au sens de Riemann*, dites encore *Riemann-intégrables.*

Définition 3.1 — *Une fonction f définie sur $[a,b]$ est dite Riemann-intégrable s'il existe $\mathrm{S}\in\mathbf{R}$, vérifiant que, pour tout $\varepsilon>0$, il existe un* nombre réel $\delta>0$ *tel que, pour toute subdivision δ-fine (pour la jauge* constante δ*), on ait*

$$|\mathrm{S}-\mathrm{S}(f,\mathcal{P})|<\varepsilon.$$

Il est bien évident que toute fonction Riemann-intégrable est intégrable, mais nous allons voir dans un instant que la réciproque est fausse : les fonctions intégrables au sens de Riemann doivent présenter une certaine uniformité dans leurs variations locales. Cependant l'intégrabilité au sens de Riemann est souvent suffisante, en particulier lorsque l'on fait du calcul intégral en relation avec le calcul différentiel (où l'on calcule des primitives de fonctions continues, qui sont Riemann-intégrables). C'est lorsque l'on a besoin d'outils d'analyse plus puissants comme de théorèmes de convergence suffisamment généraux pour traiter confortablement les procédés d'approximation classiques (séries de Fourier, par exemple) que l'intégrale de Riemann devient insuffisante. Mais donnons immédiatement l'exemple d'une fonction qui est intégrable mais pas Riemann-intégrable.

Soit $\chi : [0,1] \longrightarrow \mathbf{R}$ la fonction de Dirichlet, ou fonction indicatrice des nombres rationnels :

$$\chi(x) = 1 \text{ si } x \in \mathbf{Q} \cap [0,1], \quad \chi(x) = 0 \text{ sinon.}$$

Montrons que χ est intégrable, d'intégrale nulle. Pour cela, remarquons que $\mathrm{D} = \mathbf{Q} \cap [0,1]$ est *dénombrable,* car $\mathbf{Q}$ lui-même est dénombrable. On peut donc énumérer D en $\mathrm{D} = \{r_1, r_2, \ldots\}$. Fixons $\varepsilon > 0$ et définissons la jauge δ_ε par

$$\begin{cases} \delta_\varepsilon(x) = 1 \text{ si } x \notin \mathrm{D} \\ \delta_\varepsilon(x) = \varepsilon 2^{-j} \text{ si } x = r_j, \ j = 1, 2, \ldots \end{cases}$$

L'idée derrière la définition de cette jauge est la suivante (nous la retrouverons constamment) : si x n'est pas dans D, alors $\chi(x) = 0$ et si (I_k, x) apparaît dans une subdivision pointée, la contribution à la somme de Riemann est nulle, et ce quelle que soit la longueur de I. C'est pourquoi on pose $\delta(x) = 1$, mais on pourrait prendre n'importe quel nombre strictement positif. Si par contre x est élément de D et (I_k, x) apparaît dans une subdivision, alors $\chi(x) = 1$, et il faut compenser cela en imposant à la longueur de I_k d'être toute petite, ici suffisamment petite pour que la somme $|\mathrm{I}_1| + |\mathrm{I}_2| + \cdots$ des longueurs converge et soit inférieure à ε.

Ainsi, soit $\mathcal{P} = \{(\mathrm{I}_k, t_k), 1 \leqslant k \leqslant n\}$ une subdivision δ_ε-fine. Alors $\chi(t_k)|\mathrm{I}_k| = 0$ si $t_k \notin \mathrm{D}$ et $\chi(t_k)|\mathrm{I}_k| = \varepsilon 2^{-j}$ si $t_k = r_j$ et donc

$$0 \leqslant \mathrm{S}(\chi, \mathcal{P}) = \sum_{k=1}^{n} \chi(t_k)|\mathrm{I}_k| \leqslant \varepsilon \sum_{j=1}^{+\infty} 2^{-j} \leqslant \varepsilon\,;$$

la fonction χ est intégrable, d'intégrale nulle.

Elle n'est cependant pas intégrable au sens de Riemann. En voici succinctement la raison. Si $\mathcal{P} = \{(\mathrm{I}_k, t_k)\}$ est une subdivision δ-fine, pour une jauge *constante* δ, alors les subdivisions déduites de $\mathcal{P}$ en gardant les intervalles I_k mais en perturbant légèrement les points de marquage t_k sont encore δ-fines. Comme D et $[0,1] - \mathrm{D}$ sont denses dans $[0,1]$, on peut trouver des partitions δ-fines telles que tous les t_k soient hors de D et d'autres telles que tous soient dans D. Ainsi, pour toute constante $\delta > 0$, il existe des partitions δ-fines telles que $\mathrm{S}(f, \mathcal{P}) = 0$ et d'autres telles que $\mathrm{S}(f, \mathcal{P}) = 1 > 0$. En fait, en choisissant bien les intervalles et leurs points de marquage, on peut obtenir à la limite toute valeur entre 0 et 1. La fonction χ ne peut donc être Riemann-intégrable.

3.3.4. Deux fonctions non bornées, intégrale de Lebesgue

Pour finir, donnons deux exemples de fonctions dont on peut démontrer à la main qu'elles sont intégrables. La première est non bornée sur son intervalle de définition, la seconde est non bornée et d'intégrale non absolument convergente (*voir* le chapitre sur les intégrales impropres). Ces deux exemples montrent que l'intégrale présentée dans ce cours est une extension stricte à la fois de l'intégrale de Riemann et de l'intégrale de Lebesgue (*voir* plus loin). Les démonstrations sont laissées en exercices pour le lecteur.

Soit $f : [0,1] \longrightarrow \mathbf{R}$ la fonction définie par

$$f(0) = 0,\ f(x) = \frac{1}{\sqrt{x}} \text{ si } x \in \,]0,1].$$

Alors f est intégrable sur $[0,1]$ et

$$\int_0^1 \frac{1}{\sqrt{x}}\,\mathrm{d}x = 2.$$

Considérons pour finir la série de terme général $(-1)^n/n$. C'est une série convergente mais non absolument convergente. On peut démontrer (c'est l'objet d'un exercice très classique) qu'en réordonnant les termes de cette série on peut la faire converger vers n'importe quelle limite : elle n'est pas commutativement convergente (en fait, une série de nombres réels est commutativement convergente si et seulement si elle est absolument convergente). Construisons, sur l'intervalle $[0,1]$, une fonction en escalier dont l'intégrale soit la somme de cette série :

$$g(x) = \begin{cases} (-1)^{k+1}2^k/k \ \text{ si } \ x \in [1-2^{-(k-1)}, 1-2^{-k}[,\ k \in \mathbf{N}^* \\ 0 \ \text{ si } \ x = 1. \end{cases}.$$

Entre 0 et 1/2, g vaut $2/1$, entre 1/2 et 3/4 g vaut $-2^2/2$, etc. En fait, g est intégrable sur $[0,1]$, et

$$\int_0^1 g(x)\,\mathrm{d}x = \sum_{k=1}^{+\infty} \frac{(-1)^{k+1}}{k},$$

on a donc un exemple d'intégrale dite semi-convergente : g est intégrable, mais pas $|g|$. Cela nous amène à la définition suivante :

Définition 3.2 — *Une fonction f définie sur $\mathrm{I} = [a,b]$ est dite intégrable au sens de Lebesgue, ou* Lebesgue-intégrable, *si f* ET *$|f|$ sont intégrables.*

La classe des fonctions Lebesgue-intégrables est très importante en analyse, en particulier parce qu'elle peut se généraliser à d'autres espaces que $\mathbf{R}$: dimension plus grande, courbes, surfaces, espaces probabilisés, etc. Nous y reviendrons plus en détail dans les deuxième et troisième parties de ce livre.

3.4. Manipulations de jauges et de subdivisions

Finissons ce chapitre par deux procédures utiles dans le maniement des jauges et des subdivisions. La première consiste à couper une subdivision selon ses points de marquage. Cela est utile lorsque l'on veut par exemple mettre bout à bout deux intervalles, comme dans la formule de Chasles. La seconde consiste à forcer certains points à être de marquage, ce qui est utile lorsque la difficulté se concentre en un point particulier, de discontinuité de la fonction par exemple.

3.4.1. Coupage-collage d'une subdivision le long de ses points de marquage

Soient δ une jauge sur $[a, b]$ et $\mathcal{P} = \{([x_{k-1}, x_k], t_k), k = 1, \ldots, n\}$ une subdivision δ-fine. Formons la subdivision $\mathcal{P}^\star$ obtenue en coupant $\mathcal{P}$ en ses points de marquages :

$$\mathcal{P}^\star = \{([x_0, t_1], t_1), ([t_1, x_1], t_1), ([x_1, t_2], t_2), \ldots, ([x_{n-1}, t_n], t_n), ([t_n, b], t_n)\}.$$

Il est clair que $\mathcal{P}^\star$ est δ-fine. Mais par ailleurs, pour toute fonction f définie sur $[a, b]$, on a

$$\mathrm{S}(f, \mathcal{P}^\star) = \sum_{k=1}^{n} [f(t_k)(x_k - t_k) + f(t_k)(t_k - x_{k-1})] = \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}) = \mathrm{S}(f, \mathcal{P}),$$

et donc passer de $\mathcal{P}$ à $\mathcal{P}^\star$, et réciproquement, est sans effet, ni sur la finesse ni sur les sommes de Riemann. Remarquons que l'on aurait pu faire une procédure partielle, en ne demandant qu'à certains points t_k de devenir des extrémités.

3.4.2. Forçage des points de marquage

Considérons pour l'instant un cas simple, celui d'un point $c \in [a, b]$, et définissons la jauge $\delta_c(t)$ par

$$\begin{cases} \delta_c(c) = 1/4 \\ \delta_c(t) = |t - c|/2 \text{ si } t \neq c. \end{cases}$$

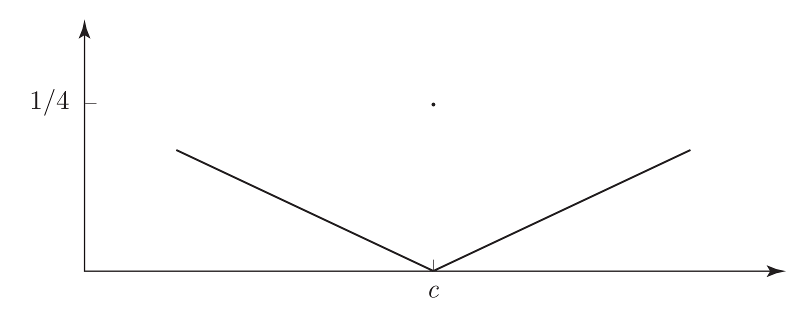


Figure 3.1 – Jauge forçant c à être point de marquage

Soit alors $\mathcal{P}$ une partition δ_c-fine et k l'indice tel que $c \in \mathrm{I}_k$. Comme $\mathcal{P}$ est δ_c-fine, on a

$$0 \leqslant |t_k - c| \leqslant \delta_c(t_k).$$

Si l'on avait $t_k \neq c$, on aurait $|t_k - c| = 2\delta_c(t_k) > 0$, contredisant l'inégalité ci-dessus. Donc $t_k = c$ et l'on a forcé c à être un point de marquage de $\mathcal{P}$.

Ce procédé peut se généraliser : soient δ une jauge et $c_1 < \cdots < c_l$ des points de $[a, b]$. Définissons alors une jauge δ' par

$$\delta'(t) = \min\{\delta(t), \delta_{c_1}(t), \ldots, \delta_{c_l}(t)\}.$$

Alors toute subdivision $\mathcal{P}$ qui est δ'-fine est aussi δ-fine et δ_{c_i}-fine, donc *les points* $c_1, \ldots, c_l$ *sont des points de marquage de* $\mathcal{P}$.

3.4.3. Application : les fonctions en escalier sont intégrables

Soient $a < c < b$, α, β deux constantes, et f la fonction en escalier valant α sur $[a, c[$, β sur $]c, b]$ et prenant une valeur arbitraire en c.

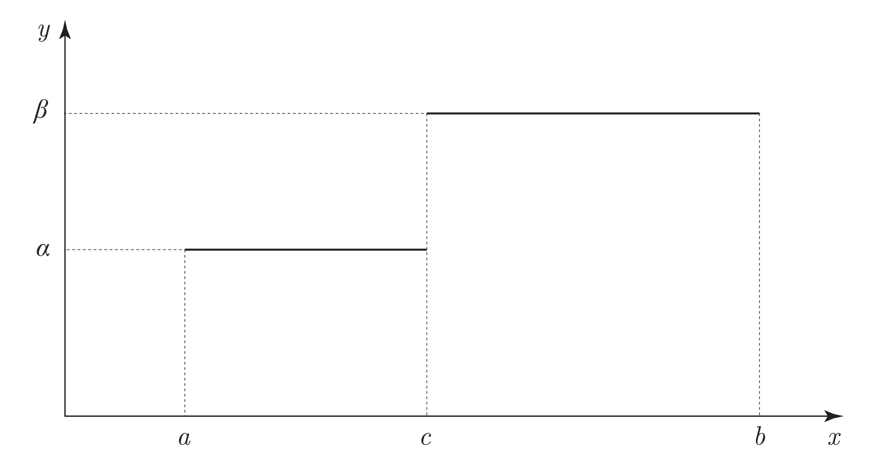


Figure 3.2 – Une fonction en escalier

Le point de discontinuité, en c, est celui qui pose problème par rapport au cas des fonctions constantes. Fixons $\varepsilon > 0$ et soit δ la jauge définie par

$$\begin{cases} \delta(x) = \dfrac{1}{2}|x - c| \text{ si } x \neq c, \\ \delta(c) = \dfrac{\varepsilon}{\max\{1, |f(c)| + |\alpha| + |\beta|\}}. \end{cases}$$

Si $\mathcal{P}$ est une subdivision δ-fine, alors c est un des points de marquage. Quitte à couper $\mathcal{P}$ le long de son point de marquage c, on peut supposer que c est sommet des deux

intervalles successifs $[x_{k-1}, c]$ et $[c, x_k]$ de $\mathcal{P}$. Mais alors les sommes de Riemann se « télescopent » sur chacun des intervalles $[a, c[$ et $[c, b]$, et l'on obtient

$$\begin{aligned}\mathrm{S}(f, \mathcal{P}) &= \alpha(x_{k-1} - a) + f(c)(x_k - x_{k-1}) + \beta(b - x_k)\\ &= \alpha(c - a) + \beta(b - c) + f(c)(x_k - x_{k-1}) + \alpha(x_{k-1} - c) + \beta(c - x_k),\end{aligned}$$

soit, par le choix de la jauge δ,

$$|\mathrm{S}(f, \mathcal{P}) - (\alpha(c - a) + \beta(b - c))| \leqslant \varepsilon.$$

Remarquons que dans l'argument ci-dessus, la valeur de f au point de discontinuité c ne joue qu'un rôle temporaire : à la limite, elle disparaît.

Définition 3.3 — *On dit qu'une fonction $f : [a, b] \longrightarrow \mathbf{R}$ est en escalier s'il existe une suite finie de points $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$, et des valeurs $\alpha_1, \ldots, \alpha_n \in \mathbf{R}$ tels que*

$$f(x) = \alpha_i \ \ si \ \ x_{i-1} < x < x_i,$$

et f prend des valeurs quelconques aux points x_i.

On aurait tout aussi bien pu dire qu'une fonction en escalier est une fonction dérivable et de dérivée nulle sur $[a, b]$ privé d'un ensemble fini de points.

Proposition 3.6 — *Soit f une fonction en escalier sur $[a, b]$. Supposons que f prenne les valeurs $\alpha_1, \ldots, \alpha_n$ sur les intervalles $\mathrm{I}_1 \leqslant \mathrm{I}_k \leqslant \cdots \leqslant \mathrm{I}_n$ (nous laissons au lecteur le soin de deviner ce que signifie cette notation saugrenue), qui forment une subdivision de $[a, b]$. Alors f est intégrable et*

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{k=1}^{n} \alpha_k |\mathrm{I}_k|.$$

Les fonctions en escalier sont très importantes en analyse, du fait de leur simplicité, et parce qu'elles permettent d'approcher, en un sens convenable, toute fonction Lebesgue-intégrable. Remarquons au passage qu'une fonction en escalier est elle-même Lebesgue-intégrable. Nous reviendrons sur ces sujets dans les derniers chapitres de ce cours. Remarquons tout de même que la fonction de Dirichlet, rencontrée un peu plus haut, est Lebesgue-intégrable : elle est en effet intégrable et positive.

3.5. Appendice : utilisation des jauges en analyse réelle

Jauges et analyse

Nous avons à plusieurs reprises utilisé des jauges provenant de définitions ponctuelles de la continuité, de la dérivabilité, etc. Elles constituent en fait le langage naturel de l'analyse, et permettent de mieux appréhender les subtiles différences entre certaines définitions.

Définition 3.4 — *On dit que f est continue sur l'intervalle $[a, b]$ si, pour tout $\varepsilon > 0$, il existe une jauge δ telle que, pour tout $x \in [a, b]$, on ait*

$$|y - x| \leqslant \delta(x) \text{ implique } |f(x) - f(y)| < \varepsilon.$$

Définition 3.5 — *On dit que f est uniformément continue sur $[a, b]$ si, dans la définition ci-dessus, on peut choisir la jauge δ constante.*

On peut également revisiter la topologie de $\mathbf{R}$:

Définition 3.6 — *On dit que* O *est un ouvert de* $\mathbf{R}$ *si, pour tout $\varepsilon > 0$, il existe une jauge δ (définie sur* O*, ou sur une partie contenant* O*,* $\mathbf{R}$ *en entier par exemple) telle que*

$$x \in \mathrm{O} \text{ implique }]x - \delta(x)/2, x + \delta(x)/2[\subset \mathrm{O}.$$

De même, la notion de dérivabilité (cf. le chapitre précédent) peut être dite en termes de jauges et, en fait, une bonne partie de l'analyse moderne, en particulier celle qui cherche à étendre les notions comme la dérivabilité, la régularité ou autres aux espaces métriques généraux est basée sur le concept de jauge (ou de recouvrement).

Nous avons vu plus haut que le lemme de Cousin était une conséquence simple du théorème des segments emboîtés, et qu'il permettait de donner une preuve simple du fait que les fermés bornés de $\mathbf{R}$ vérifient la propriété de Borel-Lebesgue. Nous allons voir qu'il permet de démontrer simplement de nombreux résultats basiques de l'analyse en une variable réelle. Commençons par donner une preuve alternative du lemme de Cousin, dans laquelle on admet que les intervalles fermés bornés vérifient la propriété de Borel-Lebesgue :

Preuve alternative du lemme de Cousin

L'intervalle $[a, b]$ étant fermé et borné, il vérifie la propriété de Borel-Lebesgue. En particulier, comme les intervalles $]t - \delta(t)/2, t + \delta(t)/2[$, $t \in [a, b]$ recouvrent $[a, b]$ et sont ouverts, on peut trouver un nombre fini d'entre eux, $]t_i - \delta(t_i)/2, t_i + \delta(t_i)/2[$, $t_1 < t_2 < \cdots < t_n$, qui recouvrent $[a, b]$.

Nous supposerons de plus que ce sous-recouvrement est minimal, dans le sens où, si on lui enlève un de ses intervalles, ce qui reste ne recouvre plus tout $[a, b]$. Construisons maintenant une subdivision pointée, avec les t_i pour points de marquages. Pour cela, il nous suffit de prendre $x_0 = a, x_n = b$, et de choisir, pour tout $1 \leqslant i < n$, un point $x_i \in [\max\{t_{i+1} - \delta(t_{i+1})/2, t_i\}, \min\{t_i + \delta(t_i)/2, t_{i+1}\}]$. En effet, le précédent intervalle est non vide du fait que les $]t_i - \delta(t_i)/2, t_i + \delta(t_i)/2[$ recouvrent, donc on peut trouver un point x_i dedans. Enfin, on a clairement $t_{i+1} \in [x_i, x_{i+1}]$ et, par minimalité du recouvrement, on a $|x_{i+1} - x_i| \leqslant \delta(t_{i+1})$. La subdivision $\{([x_i, x_{i+1}], t_{i+1}), 0 \leqslant i \leqslant n - 1\}$ est donc δ-fine.

Le théorème des valeurs intermédiaires

Soient $[a, b]$ un intervalle compact de $\mathbf{R}$ et $f : [a, b] \longrightarrow \mathbf{R}$ une fonction continue telle que pour tout $x \in [a, b]$ on ait $f(x) \neq 0$. Montrons qu'alors $f(a)f(b) > 0$, c'est-à-dire que $f(a)$ et $f(b)$ sont de même signe. En effet, comme f est continue en tout point $x \in [a, b]$ et que $f(x) \neq 0$, il existe $\delta(x) > 0$ tel que pour tout $y \in [x - \delta(x), x + \delta(x)]$ on ait $f(y)f(x) > 0$. La fonction $x \mapsto \delta(x)$ est une jauge sur $[a, b]$ et par le lemme de Cousin il existe une subdivision δ-fine $\mathcal{P} = \{(\mathrm{I}_k, t_k), k = 1, \dots, n\}$. Si l'on pose $\mathrm{I}_k = [x_{k-1}, x_k]$ on a par définition $[x_{k-1}, x_k] \subset [t_k - \delta(t_k), t_k + \delta(t_k)]$ et donc pour tout $k = 1, \dots n$ on a $f(t_k)f(x_{k-1}) > 0$ et $f(t_k)f(x_k) > 0$. Ainsi $f(x_0)$ est de même signe que $f(x_1)$, lui-même de même signe que $f(x_2)$, etc. De proche en proche, cela démontre effectivement que $f(a) = f(x_0)$ est de même signe que $f(b) = f(x_n)$.

Théorème des valeurs intermédiaires — *Soient* I *un intervalle de* $\mathbf{R}$ *et* $f : \mathrm{I} \longrightarrow \mathbf{R}$ *une fonction continue sur* $\mathbf{R}$. *Alors* $f(\mathrm{I})$ *est un intervalle.*

Démonstration — Soient $a < b$ deux éléments de I. Quitte à changer f en $-f$ on peut supposer que $f(a) \leqslant f(b)$. Le segment $[a, b]$ est entièrement inclus dans I car I est un intervalle. Nous allons commencer par démontrer que si y est un élément de $[f(a), f(b)]$, alors il existe $x \in [a, b]$ tel que $f(x) = y$. Raisonnons par l'absurde et supposons qu'un tel x n'existe pas. Alors la fonction $g(t) = f(t) - y$ ne s'annule pas sur $[a, b]$ et, d'après le raisonnement fait plus haut, on en déduit que $(f(a) - y)(f(b) - y) > 0$. Mais ceci est impossible puisque $y \in [f(a), f(b)]$ et donc nous sommes arrivés à une absurdité ; il existe donc $x \in [a, b]$ tel que $f(x) = y$.

Nous pouvons maintenant achever la démonstration du théorème en constatant que si $\alpha < \beta$ sont deux éléments de $f(\mathrm{I})$, alors ils sont de la forme $\alpha = f(a), \beta = f(b)$ et le raisonnement fait ci-dessus implique que $[\alpha, \beta] \subset \mathrm{I}$, donc que I est un intervalle. □

Le théorème de Heine

Soient $[a, b]$ un intervalle compact de $\mathbf{R}$ et f une fonction continue sur $[a, b]$. Soient par ailleurs $\varepsilon > 0$ et t un élément de $[a, b]$; comme f est continue au point t, il existe $\delta(t) > 0$ tel que si $x \in [t - \delta(t), t + \delta(t)]$, alors $|f(x) - f(t)| < \varepsilon/2$. Si x et y sont deux points de $[t - \delta(t), t + \delta(t)]$, l'inégalité triangulaire nous assure alors que $|f(x) - f(y)| < \varepsilon$. Maintenant δ définit une jauge sur $[a, b]$ et d'après le lemme de Cousin il existe une subdivision δ-fine $\mathcal{P} = \{(\mathrm{I}_k, t_k), k = 1, \dots, n\}$. Posons $\delta = \min\{\delta(t_1), \dots, \delta(t_n)\}$. Le nombre δ est strictement positif, et par ailleurs si x, y sont des éléments de $[a, b]$ tels que $|x - y| \leqslant \delta/4$, alors il existe $k \in \{1, \dots, n\}$ tel que x et y soient éléments de $[t_k - \delta(t_k), t_k + \delta(t_k)]$. Mais cela implique, par définition de la jauge δ, que $|f(x) - f(y)| < \varepsilon$. Nous venons de démontrer le

Théorème de Heine — *Une fonction continue sur un intervalle compact y est uniformément continue.*

L'image continue d'un intervalle compact est compacte

Soient $[a,b]$ un intervalle compact de $\mathbf{R}$ et f une fonction continue sur $[a,b]$. Si f n'atteignait pas sa borne supérieure (en particulier si celle-ci était égale à $+\infty$), alors pour tout $x \in [a,b]$ on pourrait trouver un autre élément de $[a,b]$, que nous noterons s_x, tel que $f(s_x) > f(x)$. Comme f est continue et que $\varepsilon_x = f(s_x) - f(x) > 0$, il existe $\delta(x) > 0$ tel que pour tout $y \in [s_x - \delta(x), s_x + \delta(x)]$ on ait

$$f(y) - f(x) \leqslant \frac{1}{2}(f(s_x) - f(x)).$$

Cela entraîne en particulier que

$$f(y) \leqslant \frac{1}{2}(f(s_x) + f(x)) < f(s_x).$$

La fonction $x \mapsto \delta(x)$ définissant une jauge sur $[a,b]$, on peut lui appliquer le lemme de Cousin et il existe donc une subdivision δ-fine $\mathcal{P} = \{(\mathrm{I}_k, t_k), k = 1, \ldots, n\}$. Soit $i \in \{1, \ldots, n\}$ un indice tel que

$$f(s_{t_i}) = \max\{f(s_{t_k}),\ k = 1, \ldots, n\}.$$

Soit maintenant y un élément quelconque de $[a,b]$. Il est élément d'un des intervalles de la subdivision $\mathcal{P}$, disons I_j et comme $\mathcal{P}$ est δ-fine on a $f(y) < f(s_{t_j}) \leqslant f(s_{t_i})$; mais cette dernière inégalité devrait alors s'appliquer au cas particulier où $y = s_{t_i}$, et cela entraîne une absurdité : $f(s_{t_i}) < f(s_{t_i})$. Ainsi f atteint-elle sa borne supérieure sur $[a,b]$ et en particulier elle est majorée et, quitte à changer f en $-f$, le même raisonnement montre que f atteint sa borne inférieure et est donc minorée. Nous avons donc démontré le théorème suivant :

Théorème 3.7 — *Soit f une fonction continue sur le segment $[a,b]$. Alors f est bornée et atteint ses bornes.*

Ainsi l'image d'un intervalle compact par une application continue est un intervalle compact. On peut facilement adapter la preuve qui précède pour démontrer, plus généralement, que l'image continue d'un compact est compacte.

L'inégalité des accroissements finis

Soient I un intervalle de $\mathbf{R}$ et f une fonction dérivable sur I. Si ε est un nombre strictement positif et $[a,b] \subset \mathrm{I}$, nous avons vu qu'il existait une jauge δ sur $[a,b]$ telle que pour tout $x \in [a,b]$ et tout u, v satisfaisant $x - \delta(x)/2 \leqslant u \leqslant x \leqslant v \leqslant x + \delta(x)/2$ on ait

$$|f(u) - f(v) - f'(x)(v-u)| \leqslant \frac{\varepsilon}{b-a}(v-u).$$

Cela entraîne en particulier que

$$|f(u) - f(v)| \leqslant \left(|f'(x)| + \frac{\varepsilon}{b-a}\right)(v-u).$$

Inégalité des accroissements finis — *Soient* I *un intervalle de* $\mathbf{R}$, f *une fonction dérivable sur* I *et* $[a, b]$ *un segment inclus dans* I. *Si* $|f'|$ *est majorée par un nombre réel* M *sur* $[a, b]$, *alors* $|f(b) - f(a)| \leqslant \mathrm{M}|b - a|$.

DÉMONSTRATION — Reprenons le cours du raisonnement fait ci-dessus. La dernière inégalité implique que

$$|f(u) - f(v)| \leqslant \Big(\mathrm{M} + \frac{\varepsilon}{b-a}\Big)(v - u).$$

La fonction δ est une jauge sur $[a, b]$, et d'après le lemme de Cousin il existe une subdivision δ-fine $\mathcal{P} = \{(\mathrm{I}_k, t_k), k = 1, \dots, n\}$. Si l'on pose $\mathrm{I}_k = [x_{k-1}, x_k]$, on tire de l'inégalité ci-dessus

$$\begin{aligned} |f(b) - f(a)| &= \Big| \sum_{k=1}^{n} f(x_k) - f(x_{k-1}) \Big| \\ &\leqslant \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \\ &\leqslant \sum_{k=1}^{n} \Big(\mathrm{M} + \frac{\varepsilon}{b-a}\Big)(x_k - x_{k-1}) \\ &= \Big(\mathrm{M} + \frac{\varepsilon}{b-a}\Big)(b - a). \end{aligned}$$

Comme cette inégalité est vraie pour tout ε, le théorème est démontré. □

This page intentionally left blank

Chapitre 4
Propriétés élémentaires de l'intégrale

Dans ce chapitre nous allons démontrer les propriétés les plus simples de l'intégrale, indispensables pour poursuivre son étude. Nous allons pour cela commencer par démontrer un critère d'intégrabilité extrêmement utile, le *critère de Cauchy*.

4.1. Critère de Cauchy, applications

L'intégrale est définie par un processus de passage à la limite. Jusqu'à maintenant, nous avons démontré des résultats d'intégrabilité pour des fonctions dont nous connaissions, *a priori*, le caractère intégrable et la valeur de l'intégrale. Pour démontrer des résultats plus généraux, il est indispensable de disposer d'un critère d'intégrabilité qui ne fasse pas appel à la valeur de l'intégrale. Pour la convergence des suites de nombres réels (et plus généralement d'éléments d'un espace complet), un tel critère existe, c'est le critère de Cauchy. Celui-ci permet souvent de démontrer qu'une suite converge sans que l'on connaisse *a priori* sa limite. Nous disposons, dans le cadre de l'intégrale, d'un outil analogue : nous l'appellerons encore critère de Cauchy.

Théorème 4.1 — *Une fonction $f : \mathrm{I} = [a, b] \longrightarrow \mathbf{R}$ est intégrable si, et seulement si, pour tout $\varepsilon > 0$, il existe une jauge δ_ε telle que, si $\mathcal{P}$ et $\mathcal{Q}$ sont deux subdivisions δ_ε-fines, alors*

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q})| \leqslant \varepsilon.$$

Démonstration — Si f est intégrable, fixons $\varepsilon > 0$ et trouvons une jauge δ telle que pour toutes subdivisions δ-fines $\mathcal{P}$ et $\mathcal{Q}$, on ait

$$\left|\mathrm{S}(f, \mathcal{P}) - \int_{\mathrm{I}} f\right| \leqslant \varepsilon/2 \quad \text{et} \quad \left|\mathrm{S}(f, \mathcal{Q}) - \int_{\mathrm{I}} f\right| \leqslant \varepsilon/2.$$

Alors $|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q})| \leqslant \varepsilon$ et le critère de Cauchy est vérifié. Réciproquement, supposons que f vérifie le critère de Cauchy. Soit (δ_n) une suite de jauges telles que si $\mathcal{P}, \mathcal{Q}$ sont δ_n fines, alors

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q})| \leqslant \frac{1}{n}.$$

Quitte à prendre le minimum des $\delta_i, i = 1, \ldots, n$, on peut supposer (δ_n) décroissante. Choisissons, pour chaque n, une subdivision $\mathcal{P}_n$ qui soit δ_n-fine. Pour tout $m > n$, $\mathcal{P}_m$ est aussi δ_n fine. De plus, $|\mathrm{S}(f, \mathcal{P}_m) - \mathrm{S}(f, \mathcal{P}_n)| \leqslant 1/n$: la suite $(\mathrm{S}(f, \mathcal{P}_n))$ est de Cauchy dans $\mathbf{R}$. Elle converge donc vers une limite, que nous noterons $\mathcal{S}$. Fixons maintenant un $\varepsilon > 0$, et prenons $n > 2/\varepsilon$, ainsi qu'une jauge $\mathcal{Q}$ qui soit δ_n-fine. Alors

$$|\mathrm{S}(f, \mathcal{Q}) - \mathcal{S}| \leqslant |\mathrm{S}(f, \mathcal{Q}) - \mathrm{S}(f, \mathcal{P}_n)| + |\mathrm{S}(f, \mathcal{P}_n) - \mathcal{S}| \leqslant 1/n + 1/n \leqslant \varepsilon,$$

ce qui montre que f est intégrable sur I. □

Comme première application du critère de Cauchy, on peut donner une caractérisation des fonctions intégrables par un procédé d'encadrement. Celui-ci permet ensuite de traiter de vastes classes de fonctions (comme les fonctions réglées), ou encore de démontrer facilement le théorème d'interversion de limite et d'intégrale pour la convergence uniforme.

Théorème 4.2 — *Une fonction f est intégrable sur* I *si et seulement si, pour tout $\varepsilon > 0$, il existe deux fonctions intégrables $f_{-\varepsilon}$ et $f_{+\varepsilon}$ telles que*

$$f_{-\varepsilon} \leqslant f \leqslant f_{+\varepsilon} \ \ \textit{et} \ \ \left| \int_{\mathrm{I}} f_{+\varepsilon} - \int_{\mathrm{I}} f_{-\varepsilon} \right| \leqslant \varepsilon.$$

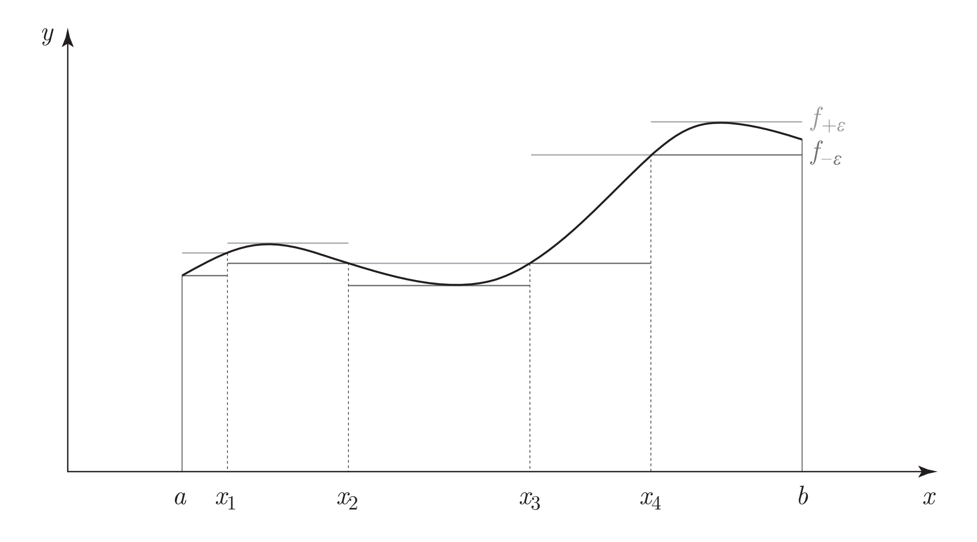


Figure 4.1 – Encadrement d'une fonction continue par deux fonctions en escalier

Démonstration — Si f est intégrable, elle vérifie clairement les conclusions du théorème, en prenant $f_{-\varepsilon} = f = f_{+\varepsilon}$. Supposons maintenant que f vérifie la condition d'encadrement ci-dessus. Si l'on se donne $\varepsilon > 0$, on trouve $f_{-\varepsilon}$ et $f_{+\varepsilon}$, ainsi que des jauges δ^+ et δ^- pour ces deux fonctions, qui sont telles que pour toute subdivision $\delta^{\pm}$-fine $\mathcal{P}$, on ait

$$\left| \mathrm{S}(f_{\pm\varepsilon}, \mathcal{P}) - \int_{\mathrm{I}} f_{\pm\varepsilon} \right| \leqslant \varepsilon.$$

On pose $\delta = \min\{\delta^+, \delta^-\}$. Pour toute subdivision pointée $\mathcal{P}$, on a

$$\mathrm{S}(f_{-\varepsilon}, \mathcal{P}) \leqslant \mathrm{S}(f, \mathcal{P}) \leqslant \mathrm{S}(f_{+\varepsilon}, \mathcal{P}).$$

Si donc $\mathcal{P}, \mathcal{Q}$ sont δ-fines, alors

$$\int_{\mathrm{I}} f_{-\varepsilon} - \varepsilon \leqslant \mathrm{S}(f, \mathcal{P}) \leqslant \int_{\mathrm{I}} f_{+\varepsilon} + \varepsilon \quad \text{et} \quad -\int_{\mathrm{I}} f_{+\varepsilon} - \varepsilon \leqslant -\mathrm{S}(f, \mathcal{Q}) \leqslant -\int_{\mathrm{I}} f_{-\varepsilon} + \varepsilon,$$

ce qui implique immédiatement, vues les hypothèses sur $f_{\pm\varepsilon}$, que

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q})| \leqslant \int_{\mathrm{I}} (f_{+\varepsilon} - f_{-\varepsilon}) + 2\varepsilon \leqslant 3\varepsilon.$$

La fonction f vérifie donc le critère de Cauchy, et est intégrable. Remarquons pour finir que la preuve montre aussi que

$$\left| \int_a^b f_{\pm\varepsilon}(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x \right| \leqslant \varepsilon. \qquad \square$$

Donnons un corollaire important de ce théorème :

Corollaire 4.3 — *Une fonction continue sur $[a, b]$ est intégrable et même Riemann-intégrable.*

Démonstration — Soit $\varepsilon > 0$. Comme f est continue sur $[a, b]$ qui est un intervalle compact, elle y est uniformément continue : il existe un nombre $\delta > 0$ tel que, si $x, y \in [a, b]$ vérifient $|x - y| < \delta$, alors $|f(x) - f(y)| < \varepsilon$. Soit n un entier tel que $1/n \leqslant \delta$, et divisons $[a, b]$ en n sous-intervalles fermés d'égale longueur $\mathrm{I}_1, \ldots, \mathrm{I}_n$. Définissons deux fonctions en escalier par

$$f_{-\varepsilon}(x) = \min\{f(x), x \in \mathrm{I}_k\} \ \text{ si } \ x \in \mathring{\mathrm{I}}_k,$$

et

$$f_{+\varepsilon}(x) = \max\{f(x), x \in \mathrm{I}_k\} \ \text{ si } \ x \in \mathring{\mathrm{I}}_k,$$

(où $\mathring{\mathrm{I}}$ est l'intérieur de l'intervalle I) et prenant la même valeur que f aux points frontières des I_k. Alors $f_{-\varepsilon}$ et $f_{+\varepsilon}$ sont des approximations uniformes de f,

car pour tout $x \in [a,b]$, on a $|f(x) - f_{\pm\varepsilon}(x)| < \varepsilon$. Par ailleurs, on a $f_{-\varepsilon} \leqslant f \leqslant f_{+\varepsilon}$, et

$$\left| \int_a^b f_{+\varepsilon}(x)\,\mathrm{d}x - \int_a^b f_{-\varepsilon}(x)\,\mathrm{d}x \right| \leqslant (b-a)\varepsilon,$$

et le théorème d'encadrement permet de conclure. □

On retrouve au passage une méthode d'intégration dérivée de celle de Riemann et due à Gaston Darboux.

Corollaire 4.4 — *Soit (f_n) une suite de fonctions intégrables sur $[a,b]$, et f une fonction définie sur $[a,b]$. Si (f_n) converge uniformément vers f, alors f est intégrable et*

$$\lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Démonstration — Pour tout n suffisamment grand, $f_n - \varepsilon$ et $f_n + \varepsilon$ encadrent f au sens du théorème d'encadrement, et la conclusion en découle. □

Le corollaire ci-dessus est souvent énoncé dans le cadre de l'intégrale de Riemann, dans lequel il est vrai et constitue, peu ou prou, le seul théorème de convergence applicable dans ce cadre. Nous verrons dans le chapitre sur les théorèmes de convergence un énoncé, le théorème de convergence dominée de Lebesgue, qui est plus général et beaucoup plus puissant.

4.2. L'intégrale comme forme linéaire

Une propriété simple à démontrer est la linéarité de l'intégrale en la fonction à intégrer.

Théorème 4.5 — *Si f et g sont des fonctions intégrables sur I, et si λ est un nombre réel, alors $f+g$ et λf sont intégrables, et l'on a les identités*

$$\int_{\mathrm{I}} (f+g) = \int_{\mathrm{I}} f + \int_{\mathrm{I}} g \quad \textit{et} \quad \int_{\mathrm{I}} \lambda f = \lambda \int_{\mathrm{I}} f.$$

Démonstration — Fixons $\varepsilon > 0$. Il existe deux jauges δ_1, δ_2 telles que pour toute subdivision $\mathcal{P}$ qui soit δ_1-fine (resp. δ_2-fine), on ait

$$\left| \mathrm{S}(f,\mathcal{P}) - \int_{\mathrm{I}} f \right| \leqslant \varepsilon \quad \text{resp.} \quad \left| \mathrm{S}(g,\mathcal{P}) - \int_{\mathrm{I}} g \right| \leqslant \varepsilon.$$

On pose alors $\delta = \min\{\delta_1, \delta_2\}$, qui est encore une jauge, et l'on note que si $\mathcal{P}$ est δ-fine, elle est aussi δ_1- et δ_2-fine. On conclut alors aisément en remarquant que pour toute subdivision $\mathcal{P}$ on a

$$\mathrm{S}(f+g,\mathcal{P}) = \mathrm{S}(f,\mathcal{P}) + \mathrm{S}(g,\mathcal{P}) \quad \text{et} \quad \mathrm{S}(\lambda f,\mathcal{P}) = \lambda \mathrm{S}(f,\mathcal{P}).$$ □

Si une fonction est positive, ses sommes de Riemann sont toutes positives et il est facile de démontrer que l'intégrale est une fonctionnelle positive :

Théorème 4.6 — *Si f est intégrable et vérifie $f \geqslant 0$, alors*

$$\int_{\mathrm{I}} f \geqslant 0.$$

En particulier, si f, g sont intégrables et vérifient $f \geqslant g$, alors

$$\int_{\mathrm{I}} f \geqslant \int_{\mathrm{I}} g.$$

On pourrait proposer une approche de l'intégrale de manière abstraite à partir des formes linéaires positives sur l'espace des fonctions continues sur $\mathbf{R}$ nulles en dehors d'un intervalle compact. On construirait ainsi, par des extensions successives de l'espace de définition de ces formes linéaires, ce que l'on appelle les mesures de Radon, dont l'intégrale de Lebesgue fournit l'exemple le plus important. On pourra à ce sujet consulter les premiers chapitres du livre de Michel Willem *Analyse harmonique réelle* paru aux éditions Hermann.

4.3. La relation de Chasles

Afin de justifier la notation avec les bornes des intervalles en haut et en bas du signe intégral, il nous faut démontrer la relation de Chasles :

Théorème 4.7 — *Soient $f : [a, b] \longrightarrow \mathbf{R}$ et $c \in \,]a, b[$. Alors f est intégrable sur $[a, b]$ si, et seulement si, ses restrictions à $[a, c]$ et $[c, b]$ le sont. Dans ce cas, on a la relation*

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f.$$

C'est pourquoi l'on note aussi l'intégrale de f sur l'intervalle $[a, b]$ sous la forme

$$\int_{[a,b]} f = \int_a^b f.$$

Démonstration — Supposons pour commencer f intégrable sur $\mathrm{I}_1 = [a, c]$ et $\mathrm{I}_2 = [c, b]$. Montrons qu'alors f est intégrable sur $\mathrm{I} = [a, b]$, et que l'on a la relation de Chasles. Pour cela, fixons $\varepsilon > 0$ et donnons-nous deux jauges δ_1 et δ_2 sur I_1 et I_2 respectivement, telles que pour toute subdivision pointée δ_i-fine $\mathcal{P}_i$ de I_i on ait

$$\left| \mathrm{S}(f, \mathcal{P}_i) - \int_{\mathrm{I}_i} f \right| \leqslant \varepsilon.$$

Définissons maintenant une jauge δ sur I en entier, qui forcera toute subdivision δ-fine à être d'une part δ_i-fine, $i = 1, 2$, et d'autre part à avoir c comme point de marquage :

$$\delta(x) = \min\left\{\delta_i(x), \frac{1}{2}|x-c|\right\}, \ \ x \in \mathrm{I}_i - \{c\}, \ \ \delta(c) = \min\left\{\delta_1(c), \delta_2(c)\right\}.$$

Par le procédé de marginalisation des points de marquage, on peut supposer que si $\mathcal{P}$ est δ-fine, c est borne d'un des intervalles de $\mathcal{P}$, et donc $\mathcal{P}$ induit sur chacun des intervalles I_i une subdivision pointée $\mathcal{P}_i$ qui est $\delta_{|\mathrm{I}_i}$-fine, donc δ_i-fine. Comme alors $\mathrm{S}(f, \mathcal{P}) = \mathrm{S}(f, \mathcal{P}_1) + \mathrm{S}(f, \mathcal{P}_2)$, il vient immédiatement que

$$\left|\mathrm{S}(f, \mathcal{P}) - \left(\int_{\mathrm{I}_1} f + \int_{\mathrm{I}_2} f\right)\right| \leqslant \left|\mathrm{S}(f, \mathcal{P}_1) - \int_{\mathrm{I}_1} f\right| + \left|\mathrm{S}(f, \mathcal{P}_2) - \int_{\mathrm{I}_2} f\right| \leqslant 2\varepsilon,$$

ce qui démontre que f est intégrable, et que la relation de Chasles est vérifiée.

Pour démontrer l'implication réciproque, le mieux est d'utiliser le critère de Cauchy. Fixons donc $\varepsilon > 0$ et soit δ une jauge sur I qui vérifie le critère de Cauchy pour ε. Si $\mathcal{P}_1$ et $\mathcal{Q}_1$ sont deux subdivisions $\delta_{|\mathrm{I}_1}$-fines, on peut, comme on l'a fait ci-dessus, leur rajouter des intervalles de sorte que c soit un point de marquage et une borne de l'intervalle de la subdivision auquel il appartient. On prolonge ensuite $\mathcal{P}_i$ et $\mathcal{Q}_i$ en deux subdivisions $\mathcal{P}$ et $\mathcal{Q}$ de I qui sont δ-fines, en utilisant les mêmes points et intervalles sur I_2. Alors

$$\mathrm{S}(f_{|\mathrm{I}_1}, \mathcal{P}_1) - \mathrm{S}(f_{|\mathrm{I}_1}, \mathcal{Q}_1) = \mathrm{S}(f, \mathcal{P}) - \mathrm{S}(f, \mathcal{Q}),$$

et comme $\mathcal{P}$ et $\mathcal{Q}$ sont δ-fine, on en déduit que

$$|\mathrm{S}(f_{|\mathrm{I}_1}, \mathcal{P}_1) - \mathrm{S}(f_{|\mathrm{I}_1}, \mathcal{Q}_1)| \leqslant \varepsilon,$$

ce qui, par le critère de Cauchy, nous permet de conclure que f est intégrable sur I_1. □

Ce théorème a pour corollaire que la restriction d'une fonction intégrable sur I à un sous-intervalle est encore intégrable. Enfin, on peut définir l'intégrale entre des bornes pas nécessairement mises dans le bon ordre, en prenant les bonnes conventions de signe. Cela sera utile lors de l'utilisation de changements de variables décroissants, et vous est déjà très familier. Pour conclure cette section, voici un autre corollaire intéressant de la relation de Chasles, qui permet souvent de simplifier les calculs :

Corollaire 4.8 — *Soient $a > 0$ et f une fonction définie sur l'intervalle $[-a, a]$. Si f est paire ou impaire alors f est intégrable sur $[-a, a]$ si et seulement si elle l'est sur $[0, a]$ ou $[-a, 0]$. Dans ce cas, on a*

$$\int_{-a}^{a} f(x)\,\mathrm{d}x = 2\int_0^a f(x)\,\mathrm{d}x \ \ \textit{si } f \textit{ est paire}$$

et

$$\int_{-a}^{a} f(x)\,\mathrm{d}x = 0 \ \ \textit{si } f \textit{ est impaire.}$$

4.4. Inégalités de Cauchy-Schwarz, Hölder et Minkowski

Soient f et g deux fonctions continues sur $[a,b]$. Les fonctions f^2, g^2 et fg sont continues et donc Lebesgue-intégrables sur $[a,b]$. Si $x \in [a,b]$, on a bien sûr

$$0 \leqslant (f(x)-g(x))^2 = f(x)^2 + g(x)^2 - 2f(x)g(x).$$

On en déduit que

$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leqslant \frac{1}{2}\Big(\int_a^b f(x)^2\,\mathrm{d}x + \int_a^b g(x)^2\,\mathrm{d}x\Big).$$

Soit maintenant $t > 0$. Le même argument que ci-dessus, utilisant l'identité remarquable $(\alpha-\beta)^2 = \alpha^2+\beta^2-2\alpha\beta$, donne, pour tout $x \in [a,b]$:

$$0 \leqslant \Big(\sqrt{t}f(x) - \sqrt{t^{-1}}g(x)\Big)^2 = tf(x)^2 + \frac{1}{t}g(x)^2 - 2f(x)g(x),$$

soit

$$2|f(x)g(x)| \leqslant tf(x)^2 + \frac{1}{t}g(x)^2.$$

Ainsi,

$$2\int_a^b |f(x)g(x)|\,\mathrm{d}x \leqslant t\int_a^b f(x)^2\,\mathrm{d}x + \frac{1}{t}\int_a^b g(x)^2\,\mathrm{d}x. \tag{4.1}$$

Si l'intégrale du carré de f est non nulle, le second membre de cette inégalité prend sa valeur minimale pour

$$t = \Big(\int_a^b g(x)^2\,\mathrm{d}x\Big)^{1/2} \Big/ \Big(\int_a^b f(x)^2\,\mathrm{d}x\Big)^{1/2},$$

ce qui, reporté dans l'inégalité (4.1), donne :

$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leqslant \Big(\int_a^b f(x)^2\,\mathrm{d}x\Big)^{1/2}\Big(\int_a^b g(x)^2\,\mathrm{d}x\Big)^{1/2}.$$

Cette inégalité est connue sous le nom *d'inégalité de Cauchy-Schwarz.* Elle est un cas particulier des inégalités de Hölder, qui se démontrent de la même manière, et qui ont pour corollaire l'inégalité de Minkowski. Pour ne pas entrer dans des détails délicats, nous énonçons toutes ces inégalités dans le cas des fonctions continues même si, nous le verrons plus loin, elles se généralisent au cas des fonctions Lebesgue-intégrables :

Théorème 4.9 — *Soient $[a,b]$ un intervalle compact et f, g des fonctions continues sur $[a,b]$.*

1. Inégalité de Cauchy-Schwarz :

$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leqslant \Big(\int_a^b f(x)^2\,\mathrm{d}x\Big)^{1/2}\Big(\int_a^b g(x)^2\,\mathrm{d}x\Big)^{1/2}.$$

2. Inégalité de Hölder : *soit $p > 1$, et q tel que $1/p + 1/q = 1$ (c.-à-d. $q = p/(p-1)$, on dit que p et q sont conjugués). Alors*

$$\int_a^b |f(x)g(x)|\,\mathrm{d}x \leqslant \left(\int_a^b f(x)^p\,\mathrm{d}x\right)^{1/p} \left(\int_a^b g(x)^q\,\mathrm{d}x\right)^{1/q}.$$

3. Inégalité de Minkowski : *soit $p > 1$, alors*

$$\left(\int_a^b |f(x)+g(x)|^p\,\mathrm{d}x\right)^{1/p} \leqslant \left(\int_a^b |f(x)|^p\,\mathrm{d}x\right)^{1/p} + \left(\int_a^b |g(x)|^p\,\mathrm{d}x\right)^{1/p}.$$

Nous laissons au lecteur le soin de démontrer l'inégalité de Hölder en s'inspirant de la preuve donnée pour l'inégalité de Cauchy-Schwarz. L'inégalité de Minkowski en est alors un corollaire.

4.5. Appendice : méthodes numériques de calcul d'intégrales

Dans l'appendice au chapitre précédent, nous avons vu une méthode naïve de calcul numérique de primitive, la méthode d'Euler. Nous allons ici considérer le problème du calcul approché d'une intégrale définie de la forme

$$\int_a^b f(x)\,\mathrm{d}x.$$

Dans de nombreuses situations où la fonction f est donnée analytiquement, le calcul exact de l'intégrale ci-dessus au moyen d'une primitive est impossible (et cela peut se démontrer rigoureusement, c'est le sujet de la théorie algébrique des équations différentielles) ; mais très souvent aussi la fonction f est elle-même uniquement connue par une approximation numérique, comme solution approchée d'une équation différentielle compliquée par exemple. Il a donc été très important, dès les débuts du calcul intégral au XVIIIe siècle, de disposer de méthodes précises et rapides pour le calcul approché d'intégrales, en particulier en astronomie.

Nous ne ferons ici qu'effleurer ce vaste sujet et renvoyons le lecteur intéressé aux quelques livres mentionnés en bibliographie. Nous nous limiterons à l'étude des méthodes de quadratures exactes pour les polynômes de degrés petits, c'est-à-dire les méthodes des rectangles, des trapèzes et de Simpson. La plupart des méthodes numériques d'intégration des fonctions d'une variable réelle fonctionnent ainsi : on se donne une subdivision de l'intervalle $[a, b]$ en n intervalles (souvent de longueurs égales) $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$, et l'on utilise la relation de Chasles pour écrire

$$\int_a^b f(x)\,\mathrm{d}x = \sum_{k=1}^{n} \int_{x_{k-1}}^{x_k} f(x)\,\mathrm{d}x.$$

Ensuite, on estime chaque intégrale sur l'intervalle $\mathrm{I}_k = [x_{k-1}, x_k]$ avec une formule faisant intervenir l'estimation de la fonction f en un certain nombre de points de I_k. On peut ne faire intervenir qu'un point, et prendre par exemple le point milieu :

$$\int_{x_{k-1}}^{x_k} f(x)\,\mathrm{d}x \approx |x_k - x_{k-1}|\, f\Big(\frac{x_k + x_{k-1}}{2}\Big),$$

et l'on obtient ce que l'on appelle couramment la *méthode des rectangles,* qui intègre de manière exacte les fonctions constantes.

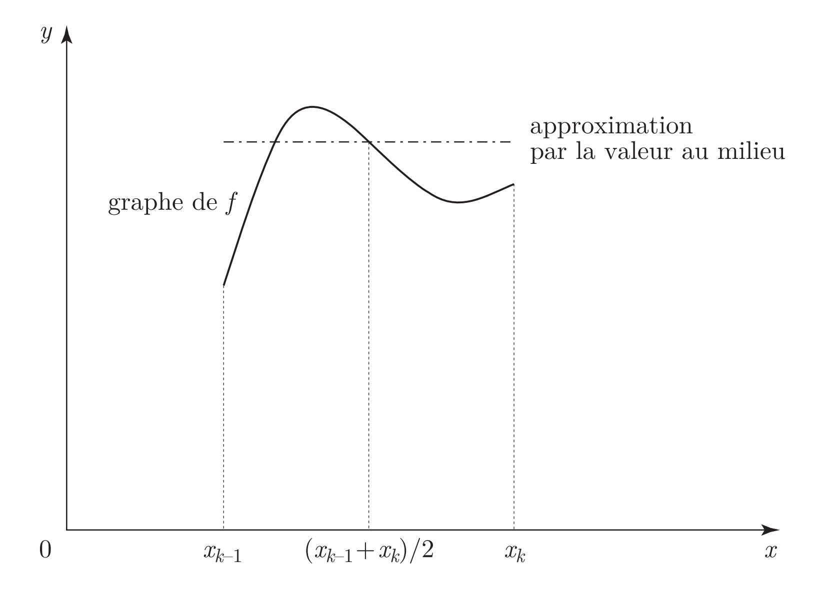


Figure 4.2 – Approximation par la méthode des rectangles

On peut aussi approcher f sur le petit intervalle $[x_{k-1}, x_k]$ par la fonction affine prenant les mêmes valeurs que f aux extrémités :

$$\int_{x_{k-1}}^{x_k} f(x)\,\mathrm{d}x \approx \frac{|x_k - x_{k-1}|}{2}\,(f(x_{k-1}) + f(x_k)),$$

ce qui constitue la *méthode des trapèzes,* qui intègre de manière exacte les fonctions affines.

Il peut être tentant d'approximer la fonction f par des polynômes de plus haut degré, afin d'obtenir des résultats plus précis. Ainsi il existe un unique polynôme de degré 2 qui prenne les mêmes valeurs que la fonctions f aux points $x_{k-1}, (x_{k-1} + x_k)/2, x_k$. En effet, un polynôme de degré 2 est donné par 3 paramètres a, b, c si on l'écrit $\mathrm{P}(x) = ax^2 + bx + c$ et imposer la valeur de P aux trois points mentionnés ci-avant revient à résoudre un système linéaire de trois équations (données par l'imposition des valeurs) à trois inconnues (les trois paramètres). On pourrait aussi, de manière plus conceptuelle, faire directement appel aux polynômes d'interpolation de Lagrange.

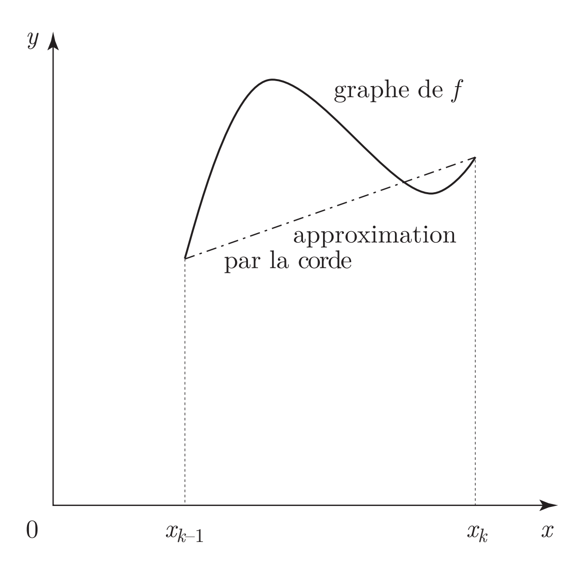


Figure 4.3 – Approximation par la méthode des trapèzes

Après un calcul élémentaire, on tire la contribution suivante pour l'intégrale de f sur le petit intervalle $[x_{k-1}, x_k]$:

$$\int_{x_{k-1}}^{x_k} f(x)\,\mathrm{d}x \approx \frac{x_k - x_{k-1}}{6}\Big(f(x_{k-1}) + 4f\Big(\frac{x_{k-1}+x_k}{2}\Big) + f(x_k)\Big).$$

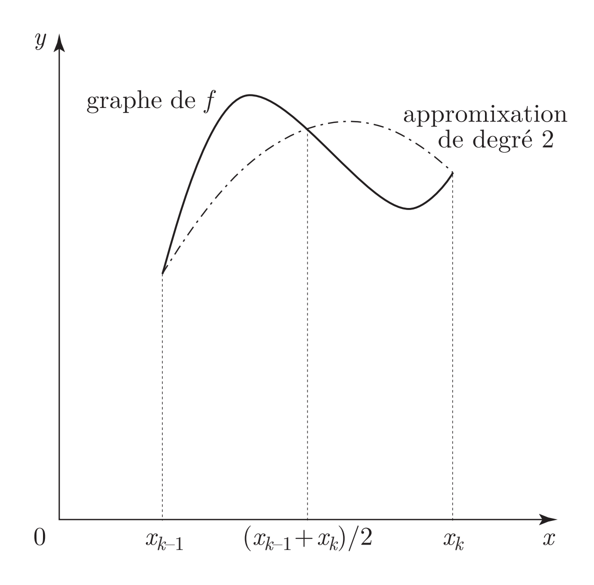


Figure 4.4 – Approximation par la méthode de Simpson

Cette méthode porte le nom de *méthode de Simpson* et est encore très utilisée dans les logiciels de calcul numérique comme MATLAB.

Signalons avant d'étudier plus précisément les méthodes des trapèzes et de Simpson qu'on peut augmenter le degré des polynômes approximant et ainsi former (*via* ce que l'on appelle les formules de quadrature de Newton-Cotes) des méthodes de calcul d'ordre supérieur, dont la précision est plus grande mais qui réclament des calculs plus laborieux. Dans la pratique, mais également pour des raisons théoriques bien précises, on s'arrête aux méthodes d'ordre 8.

Estimation de l'erreur dans la méthode des trapèzes

Nous subdivisons l'intervalle $[a, b]$ en n parties de longueurs égales $[x_0, x_1], \dots, [x_{n-1}, x_n]$ et posons $h = x_k - x_{k-1} = (b - a)/n$ ainsi que $y_k = f(x_k)$. La formule d'intégration approchée de la méthode des trapèzes peut donc s'écrire sous la forme

$$\int_a^b f(x)\,\mathrm{d}x \approx \frac{h}{2}\sum_{k=1}^{n}(y_{k-1} + y_k) = h\left(\frac{y_0}{2} + y_1 + \cdots + y_{n-1} + \frac{y_n}{2}\right)$$

et le terme d'erreur est donc

$$\mathrm{E} = \int_a^b f(x)\,\mathrm{d}x - \frac{h}{2}\sum_{k=1}^{n}(y_{k-1} + y_k) = \sum_{k=1}^{n}\left(\int_{x_{k-1}}^{x_k} f(x)\,\mathrm{d}x - \frac{h}{2}\big(f(x_{k-1}) + f(x_k)\big)\right)$$

et pour en trouver une estimation nous allons supposer que f est une fonction de classe $\mathcal{C}^2$. Pour cela, on considère la fonction auxiliaire

$$g(x) = \int_{x_{k-1}}^{x} f(t)\,\mathrm{d}t - (x - x_{k-1})\frac{f(x) + f(x_{k-1})}{2},$$

qui est elle-même de classe $\mathcal{C}^2$ et dont la dérivée est donnée par

$$g'(x) = \frac{1}{2}\big[f(x) - f(x_{k-1}) - (x - x_{k-1})f'(x)\big]$$

et la dérivée seconde par

$$g''(x) = -\frac{x - x_{k-1}}{2}f''(x).$$

Nous pouvons appliquer la formule des accroissements finis à la fonction g' sur l'intervalle $[x_{k-1}, x]$: comme $g'(x_{k-1}) = 0$ il existe $c_x \in [x_{k-1}, x]$ tel que

$$g'(x) = -\frac{(x - x_{k-1})^2}{4}f''(c_x) \tag{4.2}$$

et si l'on note m un majorant de $|f''|$ sur l'intervalle $[a, b]$ (qui existe car f'' y est continue) on obtient, pour tout $x \in [a, b]$, l'encadrement suivant :

$$-m\frac{(x - x_{k-1})^2}{2} \leqslant g'(x) \leqslant m\frac{(x - x_{k-1})^2}{2}.$$

On peut alors intégrer cette chaîne d'inégalités entre x_{k-1} et x_k, et comme $g(x_{k-1}) = 0$ on obtient

$$-m\frac{(x_k - x_{k-1})^3}{6} \leqslant g(x_k) \leqslant m\frac{(x_k - x_{k-1})^3}{6}.$$

Nous pouvons maintenant sommer les n termes de la somme, et comme nous avons pris une subdivision dont le pas est constant égal à $1/n$, nous obtenons

$$|\mathrm{E}| \leqslant \frac{(b-a)}{6n^2}m,$$

ce qui est une estimée très intéressante qui nous dit que la précision des calculs augmente comme le carré du nombre de points utilisés. Notons que l'égalité (4.2) nous informe en plus que l'erreur est de signe opposé à f'' : si f'' est partout positive, l'erreur est par défaut.

Estimation de l'erreur dans la méthode de Simpson

On applique une méthode tout à fait analogue à celle employée ci-dessus pour notre exercice de trapèze. Commençons par subdiviser l'intervalle en n parties de longueurs égales $[x_0, x_1], \ldots, [x_{n-1}, x_n]$ et posons $h = x_k - x_{k-1} = (b-a)/n$. Nous posons également $x'_k = (x_k + x_{k-1})/2$ qui est le milieu de $[x_{k-1}, x_k]$ en lequel le polynôme de degré 2 prend la même valeur que f. La formule d'approximation de l'intégrale de f sur l'intervalle $[x_{k-1}, x_k]$ donnée plus haut nous donne le terme d'erreur

$$\mathrm{E} = \int_a^b f(t)\,\mathrm{d}t - \frac{h}{3}\sum_{k=1}^{n}(f(x_{k-1}) + 4f(x'_k) + f(x_k))$$

que nous réécrivons sous la forme

$$\mathrm{E} = \sum_{k=1}^{n}\int_{x_{k-1}}^{x_k} f(t)\,\mathrm{d}t - \frac{h}{3}(f(x_{k-1}) + 4f(x'_k) + f(x_k)).$$

Nous utiliserons maintenant la fonction auxiliaire suivante :

$$g(x) = \int_{x'_k - x}^{x'_k + x} f(t)\,\mathrm{d}t - \frac{x}{3}(f(x'_k - x) + 4f(x'_k) + f(x'_k + x))$$

et supposerons que f est de classe $\mathcal{C}^4$ sur l'intervalle $[a, b]$. Il est alors aisé de démontrer que $g(0) = g'(0) = g''(0) = 0$ tandis que par ailleurs

$$g'''(x) = -\frac{x}{3}(f'''(x'_k + x) - f'''(x'_k - x)).$$

D'après la formule des accroissements finis il existe $c_x \in\,]x'_k - x, x'_k + x[$ tel que $g'''(x) = -\frac{2x^2}{3}f^{(4)}(c_x)$. On conclut alors de la même manière que pour la formule estimant l'erreur de la méthode des trapèzes et l'on trouve

$$|\mathrm{E}| \leqslant \frac{(b-a)\mathrm{M}}{1880n^4},$$

où $\mathrm{M} = \sup\{|f^{(4)}(t)|,\ t \in [a, b]\}$.

Chapitre 5
Intégrales et primitives

Dans ce chapitre nous allons montrer l'utilisation que l'on peut faire du théorème fondamental du calcul différentiel et intégral pour calculer explicitement des intégrales et des primitives. Nous allons en particulier exposer quelques méthodes pratiques de calcul de primitives, dans les cas où cela est possible.

5.1. Résultats généraux

Rappelons pour commencer le théorème fondamental (cf. § 3.2, p. 36) qui est celui sur lequel est basé tout ce chapitre.

Théorème fondamental du calcul différentiel et intégral — *Soit* $\mathrm{F} : [a, b] \longrightarrow \mathbf{R}$ *une fonction continue sur* $[a, b]$*, dérivable sur* $]a, b[$ *et admettant une dérivée à droite en* a *(resp. à gauche en* b*). Notons* f *la dérivée de* F*. Alors* f *est intégrable et l'on a*

$$\int_a^b f = \mathrm{F}(b) - \mathrm{F}(a).$$

Passons maintenant brièvement en revue les énoncés qui vous sont familiers depuis la Terminale, et commençons par introduire les *intégrales indéfinies.* Soit f une fonction intégrable sur $[a, b]$; on sait, comme l'intégrale a la propriété de restriction aux sous-intervalles, que f est intégrable sur $[a, x]$ pour tout $x \in [a, b]$, et l'on peut donc définir

$$\psi_f(x) = \int_a^x f, \ \ x \in [a, b].$$

La relation de Chasles et le théorème fondamental impliquent immédiatement le théorème suivant, base du calcul intégral au lycée :

Théorème 5.1 — *Toute fonction continue* f *admet une primitive, égale à l'addition d'une constante près à* ψ_f*.*

On peut résumer tout cela de manière condensée (mais peu rigoureuse) :

– si f est dérivable, on a
$$\int_a^x f'(t)\,\mathrm{d}t = f(x) + \text{constante}\,;$$

– si f est continue,
$$\left(\int_a^x f(t)\,\mathrm{d}t\right)' = f(x).$$

Le théorème fondamental a pour corollaire immédiat la *formule d'intégration par parties* :

Théorème 5.2 — *Soient* F *et* G *deux fonctions dérivables sur* $[a,b]$*. Alors* $\mathrm{F'G}$ *est intégrable si et seulement si* $\mathrm{FG'}$ *l'est. Dans ce cas, on a la formule*
$$\int_a^b \mathrm{F'G} = [\mathrm{FG}]_a^b - \int_a^b \mathrm{FG'}.$$

Démonstration — Comme F et G sont dérivables, FG l'est aussi et
$$(\mathrm{FG})' = \mathrm{F'G} + \mathrm{FG'}.$$

Le théorème fondamental dit par ailleurs que $(\mathrm{FG})'$ est intégrable, et le théorème s'en déduit facilement. □

Le *théorème de changement de variable*

Théorème 5.3 — *Soient* g *une application dérivable sur l'intervalle compact* $[a,b]$*, et* f *une fonction admettant une primitive sur l'intervalle* $g([a,b])$*. Alors*
$$\int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x = \int_a^b f(g(t))g'(t)\,\mathrm{d}t.$$

Démonstration — Soit F une primitive de f. Par hypothèse $\mathrm{F}\circ g$ est définie sur $[a,b]$ de dérivée donnée par $(\mathrm{F}\circ g)'(t) = f(g(t))g'(t)$. Le théorème fondamental nous dit alors que
$$\int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x = \mathrm{F}(g(b)) - \mathrm{F}(g(a)) = \int_a^b f(g(t))g'(t)\,\mathrm{d}t,$$
ce qui démontre le théorème. □

Remarquons qu'on ne suppose pas que g est une bijection. Cependant, l'intervalle $g([a,b])$ pourrait être plus grand que $[g(a), g(b)]$, et l'hypothèse que f admet une primitive sur $g([a,b])$ est donc relativement forte.

On énonce souvent le théorème de changement de variable sous la forme suivante, qui est une conséquence simple du théorème énoncé ci-dessus :

Corollaire 5.4 — *Soient* $[a, b]$ *et* $[c, d]$ *deux intervalles compacts,* $g : [a, b] \longrightarrow [c, d]$ *un difféomorphisme et* f *une application admettant une primitive sur* $[c, d]$*. Alors*

$$\int_c^d f(y)\,\mathrm{d}y = \int_a^b f(g(x))\,|g'(x)|\,\mathrm{d}x.$$

Précisons que l'hypothèse faite sur g dans le corollaire signifie que g est une bijection continue de $[a, b]$ sur $[c, d]$ et que g est dérivable sur $[a, c]$, de dérivée partout non-nulle. La formule ci-dessus porte aussi le nom de *formule du jacobien.*

Finissons ce chapitre par un théorème utile, la *formule de la moyenne* pour les intégrales. Ce résultat ne concerne pas les liens entre intégration et primitives, mais il est proche de la formule des accroissements finis, c'est pourquoi nous l'énonçons ici.

Théorème 5.5 — *Soient* f, g *deux fonctions définies sur* $[a, b]$*. On suppose* f, g *et* fg *intégrables,* g *positive, et* f *bornée. Alors*

$$\inf\big\{f(x), x \in [a, b]\big\} \int_a^b g(x)\,\mathrm{d}x \leqslant \int_a^b f(x)g(x)\,\mathrm{d}x \leqslant \sup\big\{f(x), x \in [a, b]\big\} \int_a^b g(x)\,\mathrm{d}x.$$

Si de plus f *est continue, alors il existe un point* $c \in [a, b]$ *tel que*

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(c) \int_a^b g(x)\,\mathrm{d}x.$$

Remarquons, avant de donner la démonstration, que l'hypothèse d'intégrabilité faite sur le produit fg est indispensable : il existe des fonctions f, g intégrables telles que fg ne le soit pas.

Démonstration — Soit m (resp. M) la borne inférieure (resp. supérieure) de f sur $[a, b]$. On a, pour tout $x \in [a, b]$, l'encadrement

$$mg(x) \leqslant f(x)g(x) \leqslant \mathrm{M}g(x),$$

et la première partie du théorème découle de la positivité de l'intégrale. Quant à la deuxième, elle découle du théorème des valeurs intermédiaires. □

Corollaire 5.6 — *Si* f *est continue sur* $[a, b]$*, alors il existe* $c \in \,]a, b[$ *telle que*

$$\int_a^b f(x)\,\mathrm{d}x = (b - a)f(c).$$

5.2. Quelques méthodes pratiques de calcul d'intégrales et de primitives

Le théorème fondamental et ses corollaires donnent la manière la plus efficace pour calculer explicitement des intégrales et des primitives de fonctions construites à partir de fonctions usuelles. Nous allons, dans cette section, rappeler les primitives des fonctions usuelles, puis donner quelques exemples de méthode de calcul d'intégrales. Une fois ces quelques méthodes apprises, il faut les pratiquer intensivement en résolvant de nombreux exercices. Quelques-uns sont donnés à la fin de ce livre, mais le lecteur est invité à en chercher de nombreux autres.

5.2.1. Primitives usuelles

La liste de la table 5.1 devrait être connue par cœur par tout étudiant sérieux. Pour fixer les notations, nous noterons la primitive F d'une fonction continue f, là où elle est définie, sous la forme

$$\mathrm{F} = \int f(x)\,\mathrm{d}x + \mathrm{C},$$

où C désigne une constante quelconque.

Ces formules se déduisent facilement des formules de trigonométrie (usuelle ou hyperbolique) ainsi que des formules donnant les dérivées des fonctions usuelles. De manière générale, le calcul de primitives ou d'intégrales, hormis quelques méthodes connues pour des cas particuliers, demande du nez, et donc beaucoup d'entraînement. Il est à noter qu'il n'existe pas d'algorithme général permettant de donner la primitive explicite d'une fonction définie par des fonctions usuelles. Il existe même des fonctions très simples (comme $\exp(-x^2)$) dont la primitive n'admet pas d'expression algébrique en termes de fonction usuelles. C'est pourquoi, quand on rencontre un problème de calcul explicite d'intégrale, on fait souvent appel, suivant le type d'étude entreprise, aux méthodes numériques, ou bien à des méthodes avancées d'algèbre, de géométrie ou d'analyse.

5.2.2. Intégration des fractions rationnelles

En algèbre, on définit une fraction rationnelle à coefficients dans un corps $\mathbf{K}$ comme un élément du corps des fractions de l'anneau des polynômes $\mathbf{K}[t]$. Pour nous, une fraction rationnelle sera simplement une fonction définie comme le quotient P/Q de deux polynômes P et Q à coefficients complexes. Cette fonction est définie sur $\mathbf{C}$ privé des racines de Q, que l'on appelle les *pôles* de la fraction rationnelle P/Q. Nous admettrons le théorème suivant, conséquence du fait que $\mathbf{C}$ est *algébriquement clos :* tout polynôme non constant à coefficients complexes admet au moins une racine dans $\mathbf{C}$.

Table 5.1 – Liste de primitives usuelles

$\int x^\alpha \,\mathrm{d}x = \frac{1}{\alpha+1}x^{\alpha+1} + \mathrm{C},\ \alpha \neq -1$	$\int \frac{1}{x}\,\mathrm{d}x = \mathrm{Log}\,\lvert x\rvert + \mathrm{C}$
$\int \mathrm{e}^{(a+ib)x}\,\mathrm{d}x = \frac{1}{a+ib}\mathrm{e}^{(a+ib)x} + \mathrm{C},\ a+ib \neq 0$	$\int \sin x\,\mathrm{d}x = -\cos x + \mathrm{C}$
$\int \cos x\,\mathrm{d}x = \sin x + \mathrm{C}$	$\int \tan x\,\mathrm{d}x = -\mathrm{Log}\,\lvert\cos x\rvert + \mathrm{C}$
$\int \frac{1}{\sin x}\,\mathrm{d}x = \mathrm{Log}\left\lvert\tan\frac{x}{2}\right\rvert + \mathrm{C}$	$\int \frac{1}{\cos x}\,\mathrm{d}x = \mathrm{Log}\left\lvert\tan\left(\frac{\pi}{4}+\frac{x}{2}\right)\right\rvert + \mathrm{C}$
$\int \mathrm{cotan}\,x\,\mathrm{d}x = \mathrm{Log}\,\lvert\sin x\rvert + \mathrm{C}$	$\int \frac{1}{\cos^2 x}\,\mathrm{d}x = \tan x + \mathrm{C}$
$\int \frac{1}{\sin^2 x} = -\mathrm{cotan}\,x + \mathrm{C}$	$\int \frac{1}{\sin x\cos x}\,\mathrm{d}x = \mathrm{Log}\,\lvert\tan x\rvert + \mathrm{C}$
$\int \mathrm{sh}\,x\,\mathrm{d}x = \mathrm{ch}\,x + \mathrm{C}$	$\int \mathrm{ch}\,x\,\mathrm{d}x = \mathrm{sh}\,x + \mathrm{C}$
$\int \mathrm{th}\,x\,\mathrm{d}x = \mathrm{Log}\,\mathrm{ch}\,x + \mathrm{C}$	$\int \frac{1}{\mathrm{sh}\,x}\,\mathrm{d}x = \mathrm{Log}\left\lvert\mathrm{th}\,\frac{x}{2}\right\rvert + \mathrm{C}$
$\int \frac{1}{\mathrm{ch}\,x}\,\mathrm{d}x = 2\arctan(\mathrm{e}^x) + \mathrm{C}$	$\int \frac{1}{\mathrm{th}\,x}\,\mathrm{d}x = \mathrm{Log}\,\lvert\mathrm{sh}\,x\rvert + \mathrm{C}$
$\int \frac{1}{\mathrm{ch}^2 x}\,\mathrm{d}x = \mathrm{th}\,x + \mathrm{C}$	$\int \frac{1}{\mathrm{sh}^2 x}\,\mathrm{d}x = -\coth x + \mathrm{C}$
$\int \frac{1}{\mathrm{sh}\,x\,\mathrm{ch}\,x}\,\mathrm{d}x = \mathrm{Log}\,\lvert\mathrm{th}\,x\rvert + \mathrm{C}$	$\int \frac{1}{x^2+a^2}\,\mathrm{d}x = \frac{1}{a}\arctan\frac{x}{a} + \mathrm{C}$
$\int \frac{1}{x^2-a^2}\,\mathrm{d}x = \frac{1}{2a}\mathrm{Log}\left\lvert\frac{x-a}{x+a}\right\rvert + \mathrm{C}$	$\int \frac{1}{\sqrt{a^2-x^2}}\,\mathrm{d}x = \arcsin\frac{x}{\lvert a\rvert} + \mathrm{C}$
$\int \frac{1}{\sqrt{x^2+a^2}}\,\mathrm{d}x = \mathrm{Log}(x+\sqrt{x^2+a^2}) + \mathrm{C}$	$\int \frac{1}{\sqrt{x^2-a^2}}\,\mathrm{d}x = \mathrm{Log}\,\lvert x+\sqrt{x^2-a^2}\rvert + \mathrm{C}$
$\int \frac{1}{(x^2+a)^{3/2}}\,\mathrm{d}x = \frac{x}{a\sqrt{x^2+a}} + \mathrm{C}$	$\int \frac{1}{(a-x^2)^{3/2}}\,\mathrm{d}x = \frac{x}{a\sqrt{a-x^2}} + \mathrm{C},\ a \neq 0$

Théorème 5.7 — *Soient* P *et* Q *deux polynômes à coefficients complexes sans racine commune. Soient* $a_1, \dots, a_r$ *les pôles de* P/Q. *Alors il existe un polynôme* E, *des nombres entiers* $\alpha_i \in \mathbf{N}^*, 1 \leqslant i \leqslant r$ *(*α_i *est appelé multiplicité du pôle* a_i*), et des nombres complexes* $\mathrm{A}_{j,k}, 1 \leqslant k \leqslant r, 1 \leqslant j \leqslant \alpha_k$, *tels que*

$$\frac{\mathrm{P}(x)}{\mathrm{Q}(x)} = \mathrm{E}(x) + \sum_{k=1}^{r} \sum_{j=1}^{\alpha_k} \frac{\mathrm{A}_{j,k}}{(x - a_k)^j}.$$

La décomposition ci-dessus s'appelle *décomposition de* P/Q *en éléments simples.* Le problème du calcul pratique de cette décomposition est délicat, et nous ne traiterons ici que deux exemples simples. Le lecteur est invité à faire de nombreux exercices pour se familiariser avec cette pratique. Avant de passer aux exemples, notons que comme il n'existe pas d'algorithme permettant de calculer les racines d'un polynôme général, le théorème ci-dessus est d'une portée limitée dans le calcul explicite de primitives ou d'intégrales. Par contre, nous verrons aussi que l'on ramène souvent un calcul d'intégrale, *via* des changements de variables ou tout autre méthode, à un calcul d'intégrale de fraction rationnelle.

Premier exemple

Soit à calculer l'intégrale suivante, pour $1 < a < x$:

$$\mathrm{J} = \int_a^x \frac{1}{t^3(1+t^3)}\,\mathrm{d}t.$$

Il faut commencer par vérifier que la fonction sous le signe intégrale est bien intégrable. Pour cela, remarquons que le dénominateur ne s'annule qu'en 0 et -1, donc la fonction à intégrer est intégrable sur $[a, x] \subset \mathbf{R} - \{0, -1\}$, car continue. Avant de partir tête baissée dans le calcul du développement en éléments simples de la fonction, cherchons à la simplifier. Ainsi, on remarque que

$$\frac{1}{t^3(1+t^3)} = \frac{1}{t^3} - \frac{1}{1+t^3}.$$

Le premier terme du second membre est facile à intégrer, si l'on se rappelle la règle de dérivation des puissances :

$$\int_a^x \frac{1}{t^3}\,\mathrm{d}t = \left[-\frac{1}{2}\frac{1}{t^2}\right]_a^x,$$

où la notation $[f(t)]_a^b$ signifie $f(b) - f(a)$, comme à l'accoutumée. Il nous reste donc à calculer l'intégrale du second terme. Remarquons que $1 + t^3$ a pour racines complexes $-1, \mathrm{e}^{i\pi/3}$ et $\mathrm{e}^{-i\pi/3}$. Ainsi

$$1 + t^3 = (t+1)(t - \mathrm{e}^{i\pi/3})(t - \mathrm{e}^{-i\pi/3}).$$

Les pôles de $(1 + t^3)^{-1}$ sont de multiplicité 1 (et l'on parle alors de pôles *simples*), donc la décomposition en éléments simples prend la forme

$$\frac{1}{1+t^3} = \frac{a}{t+1} + \frac{b}{t - \mathrm{e}^{i\pi/3}} + \frac{c}{t - \mathrm{e}^{-i\pi/3}}.$$

On détermine les trois nombres complexes a, b, c en multipliant successivement l'identité ci-dessus par les monômes $t+1$, $t - \mathrm{e}^{i\pi/3}$ et $t - \mathrm{e}^{-i\pi/3}$, et en faisant tendre la variable vers chacune des racines correspondantes de $1+t^3$. Commençons par déterminer a. L'identité ci-dessus implique que

$$\frac{t+1}{1+t^3} = \frac{1}{(t-\mathrm{e}^{i\pi/3})(t-\mathrm{e}^{-i\pi/3})} = a + (t+1)\left(\frac{b}{t-\mathrm{e}^{i\pi/3}} + \frac{c}{t-\mathrm{e}^{-i\pi/3}}\right),$$

et l'on peut faire tendre t vers -1 dans cette dernière égalité, pour obtenir

$$\frac{1}{(-1-\mathrm{e}^{i\pi/3})(-1-\mathrm{e}^{-i\pi/3})} = a,$$

soit $a = 1/3$. De même, l'identité plus haut, multipliée par $t - \mathrm{e}^{i\pi/3}$ donne

$$\frac{1}{(t+1)(t-\mathrm{e}^{-i\pi/3})} = b + (t-\mathrm{e}^{i\pi/3})\left(\frac{a}{t+1} + \frac{b}{t-\mathrm{e}^{-i\pi/3}}\right),$$

et, si l'on fait tendre t vers $\mathrm{e}^{i\pi/3}$, on trouve enfin

$$\frac{1}{(1+\mathrm{e}^{i\pi/3})(\mathrm{e}^{i\pi/3}-\mathrm{e}^{-i\pi/3})} = b,$$

soit

$$b = \frac{1}{i\sqrt{3}(1+\mathrm{e}^{i\pi/3})}.$$

On calcule de même c et l'on trouve finalement

$$\frac{1}{1+t^3} = \frac{1/3}{1+t} + \frac{1}{i\sqrt{3}(1+\mathrm{e}^{i\pi/3})}\frac{1}{(t-\mathrm{e}^{i\pi/3})} - \frac{1}{i\sqrt{3}(1+\mathrm{e}^{-i\pi/3})}\frac{1}{(t-\mathrm{e}^{-i\pi/3})}.$$

Le premier terme est facile à intégrer, mais les deux suivants, qui semblent tout aussi faciles, comportent des nombres complexes. Remarquons cependant que les deux derniers termes sont *conjugués complexes*. Ceci est un fait général, dû au caractère réel du polynôme $1+t^3$: les racines complexes d'un polynômes réel sont deux à deux conjuguées. Il en découle que dans la décomposition en éléments simples, on peut ajouter deux à deux les termes correspondants à un couple de pôles conjugués, et obtenir finalement une expression réelle ! Additionnons donc les deux derniers termes de l'identité ci-dessus :

$$\frac{1}{i\sqrt{3}(1+\mathrm{e}^{i\pi/3})}\frac{1}{(t-\mathrm{e}^{i\pi/3})} - \frac{1}{i\sqrt{3}(1+\mathrm{e}^{-i\pi/3})}\frac{1}{(t-\mathrm{e}^{-i\pi/3})} = \frac{1}{3}\frac{2-t}{t^2-t+1}.$$

On peut finalement écrire J sous la forme :

$$\mathrm{J} = \int_a^x \frac{1}{t^3(1+t^3)}\,\mathrm{d}t = \int_a^x \frac{1}{t^3}\,\mathrm{d}t - \frac{1}{3}\int_a^x \frac{1}{1+t}\,\mathrm{d}t - \frac{1}{3}\int_a^x \frac{2-t}{t^2-t+1}\,\mathrm{d}t.$$

Les deux premiers termes du second membre s'intègrent facilement. Nous allons voir qu'il en est de même du troisième. Pour cela, la première chose à faire est de faire

apparaître la dérivée du dénominateur au numérateur, et de compenser habilement le reste pour obtenir la somme d'une dérivée logarithmique et d'une dérivée de fonction circulaire inverse (ici l'arc-tangente). Voilà comment l'on procède : remarquons tout d'abord que

$$\frac{\mathrm{d}}{\mathrm{d}t}(t^2 - t + 1) = 2t - 1.$$

Or on ne dispose au numérateur que de $-t = -1/2(2t)$. Faisons donc apparaître $2t-1$ au numérateur :

$$\begin{aligned}\frac{2-t}{t^2-t+1} &= \frac{-1/2(2t)+2}{t^2-t+1} \\ &= \frac{-1/2(2t) - 1/2(-1) + 2 + 1/2(-1)}{t^2-t+1} \\ &= -\frac{1}{2}\frac{2t-1}{t^2-t+1} + \frac{3}{2}\frac{1}{t^2-t+1}.\end{aligned}$$

Le premier terme du dernier membre s'intègre aisément, car on y reconnaît une dérivée logarithmique. Pour le second terme, de la forme

$$\frac{1}{t^2+\alpha t+\beta},$$

on s'arrange pour le mettre sous la forme

$$\frac{1}{(t+\eta)^2+1}.$$

Remarquons que $t^2 - t$ est le début du développement de $(t - 1/2)^2$. Ainsi

$$t^2 - t + 1 = (t - 1/2)^2 + 3/4,$$

et

$$\frac{2-t}{t^2-t+1} = -\frac{1}{2}\frac{2t-1}{t^2-t+1} + \frac{2}{\frac{4}{3}\left(t-\frac{1}{2}\right)^2+1}.$$

On peut alors se reporter à la liste des intégrales usuelles, et obtenir :

$$\mathrm{J} = \left[-\frac{1}{2}\frac{1}{t^2} - \frac{1}{3}\operatorname{Log}|1+t| + \frac{1}{6}\operatorname{Log}|t^2-t+1| - \frac{\sqrt{3}}{3}\arctan\frac{2}{\sqrt{3}}\left(t-\frac{1}{2}\right)\right]_a^x.$$

Ouf ! Comme on peut le voir, les calculs sont fastidieux, mais, dès que l'on connaît les pôles de la fraction rationnelle (c.-à-d. les racines du dénominateur), ils deviennent automatiques, et l'on pourrait écrire un programme permettant de calculer l'intégrale.

Deuxième exemple

Essayons maintenant de calculer, avec les mêmes hypothèses sur a et x que plus haut, l'intégrale

$$J = \int_a^x \frac{3t^3 + 10t^2 - 2t}{(t^2-1)^2}\,dt = \int_a^x \frac{P(t)}{Q(t)}\,dt.$$

Remarquons tout d'abord que Q est de degré plus grand que P. Cela implique que le quotient de P par Q dans la division euclidienne des polynômes est 0. C'est la partie entière de P/Q, notée E dans le théorème de décomposition en éléments simples. Celui-ci nous indique donc qu'il existe des nombres complexes a, b, c, d tels que

$$\frac{P(t)}{Q(t)} = \frac{a}{(t-1)^2} + \frac{b}{t-1} + \frac{c}{(t+1)^2} + \frac{d}{t+1},$$

puisque $(t^2-1)^2$ admet -1 et 1 pour racines doubles (c.-à-d. -1 et $+1$ sont des pôles d'ordre 2 de P/Q). Remarquons que ni -1 ni 1 ne sont des racines de P. On calcule a en multipliant l'identité ci-dessus par $(t-1)^2$, ce qui donne

$$\frac{3t^3 + 10t - 2t}{(t+1)^2} = a + (t-1)\left[b + \frac{c(t-1)}{(t+1)^2} + \frac{d(t-1)}{t+1}\right].$$

En faisant $t = 1$ on obtient $a = 11/4$. De même, on obtient c en multipliant par $(t+1)^2$ et en faisant $t = -1$ dans l'égalité obtenue : $c = 9/4$. Pour trouver ensuite b et d, on affecte à t diverses valeurs, comme par exemple $t = 0$, ce qui donne $a - b + c + d = 0$. On peut aussi multiplier l'identité plus haut par t, et faire tendre t vers $+\infty$, ce qui donne $b + d = 3$. Toute autre cuisine est bien sûr possible et finalement on obtient un système de deux équations linéaires dont b et d sont solutions :

$$b - d = 5 \text{ et } b + d = 3,$$

ce qui donne $b = 4$ et $d = -1$. On peut donc maintenant calculer l'intégrale J :

$$\begin{aligned} J &= \frac{11}{4}\int_a^x \frac{1}{(t-1)^2}\,dt + \frac{9}{4}\int_a^x \frac{1}{(t+1)^2}\,dt + 4\int_a^x \frac{1}{t-1}\,dt - \int_a^x \frac{1}{t+1}\,dt \\ &= \left[-\frac{1}{2}\frac{10t+1}{t^2-1} + \operatorname{Log}\frac{(t-1)^4}{|t+1|}\right]_a^x. \end{aligned}$$

5.2.3. Applications de la formule d'intégration par parties

Premier exemple

Soit n un entier naturel, posons

$$I_n = \int_0^{\pi/2} \sin^n t\,dt,$$

intégrale portant le nom du mathématicien anglais John Wallis (1616–1703). Pour calculer une telle intégrale, une méthode courante consiste à trouver, par une ou plusieurs intégrations par parties, une relation de récurrence satisfaite par la suite (I_n).

Si $n \geqslant 2$, on a $\sin^n t = \sin^{n-1} t \sin t$, et l'on peut appliquer la formule d'intégration par parties à cette décomposition :

$$\begin{aligned}
\mathrm{I}_n &= \int_0^{\pi/2} \sin^n t \,\mathrm{d}t = \int_0^{\pi/2} \sin t \sin^{n-1} t \,\mathrm{d}t \\
&= \left[-\cos t \sin^{n-1} t\right]_0^{\pi/2} - \int_0^{\pi/2} (-\cos t)(n-1)\cos t \sin^{n-2} t \,\mathrm{d}t \\
&= (n-1)\int_0^{\pi/2} (1-\sin^2 t)\sin^{n-2} t \,\mathrm{d}t \\
&= (n-1)\mathrm{I}_{n-2} - (n-1)\mathrm{I}_n,
\end{aligned}$$

ce qui donne, pour tout $n \geqslant 2$:

$$\mathrm{I}_n = \frac{n-1}{n}\mathrm{I}_{n-2}.$$

Deux cas se présentent alors : si $n = 2p$ est pair, on obtient

$$\begin{aligned}
\mathrm{I}_{2p} &= \frac{(2p-1).(2p-3)\ldots 3.1}{(2p).(2p-2)\ldots 4.2}\,\mathrm{I}_0 \\
&= \frac{\pi}{2}\frac{(2p)!}{(2^p p!)^2},
\end{aligned}$$

où, pour passer de la première à la seconde ligne, on a multiplié numérateur et dénominateur par $2p.(2p-2)\ldots 2$ de manière à faire apparaître $(2p)!$ au numérateur et $(2p.(2p-2)\ldots 2)^2 = (2^p p!)^2$ au dénominateur. Le facteur $\pi/2$ provient du terme I_0.

De même, si $n = 2p+1$ est impair, on obtient

$$\mathrm{I}_{2p+1} = \frac{(2^p p!)^2}{(2p+1)!}.$$

On remarque qu'ici le terme I_1 donne un facteur égal à 1. Ainsi les intégrales de Wallis prennent alternativement une valeur rationnelle et une valeur irrationnelle (et même transcendante).

Deuxième exemple

Soit $n \geqslant 1$ un entier naturel, on cherche à calculer une primitive F_n de $f_n(x) = x^n \cos x$. La fonction f est déjà écrite comme produit de deux fonctions, qu'en plus l'on sait intégrer. Intégrons par parties :

$$\begin{aligned}
\int x^n \cos x \,\mathrm{d}x &= [x^n \sin x] - \int n x^{n-1} \sin x \,\mathrm{d}x \\
&= x^n \sin x - n\left([x^{n-1}(-\cos x)] - \int (n-1)x^{n-2}(-\cos x)\,\mathrm{d}x\right) \\
&= x^n \sin x + n x^{n-1}\cos x - n(n-1)\mathrm{F}_{n-2}(x),
\end{aligned}$$

relation de récurrence qui permet de calculer F_n pour tout n.

5.2.4. Applications de la formule de changement de variable

Commençons par rappeler un moyen mnémotechnique utile dans l'utilisation de la formule de changement de variable. Si l'on pose $x = \varphi(t)$, alors on obtient $\mathrm{d}x$ par la formule $\mathrm{d}x = \varphi'(t)\,\mathrm{d}t$.

Un premier exemple amusant

Essayons de déterminer une primitive de

$$\frac{1}{\sqrt{1+x}+\sqrt[3]{1+x}}.$$

On remarque que l'on a les mêmes quantités sous les radicaux. On va donc essayer de s'en débarrasser en exprimant ces deux radicaux comme puissances d'une même quantité. Le ppcm de 2 (pour la racine carrée) et 3 est 6. Ainsi on a

$$(\sqrt[6]{1+x})^2 = \sqrt[3]{1+x} \text{ et } (\sqrt[6]{1+x})^3 = \sqrt{1+x}.$$

On s'empresse donc de faire le changement de variable $y = \sqrt[6]{1+x} = (1+x)^{1/6}$. Comme $y^6 = 1+x$, il vient $\mathrm{d}x = 6y^5\,\mathrm{d}y$, et donc

$$\int \frac{1}{\sqrt{1+x}+\sqrt[3]{1+x}}\,\mathrm{d}x = \int \frac{6y^5}{y^3+y^2}\,\mathrm{d}y,$$

et cette dernière intégrale se calcule facilement. On commence par effectuer une division euclidienne de y^5 par y^3+y^2, ce qui donne

$$y^5 = (y^2-y+1)(y^3+y^2) - y^2,$$

soit

$$\frac{y^5}{y^3+y^2} = y^2 - y + 1 - \frac{1}{y+1},$$

et il sort alors de tout ça le merveilleux résultat suivant :

$$\int \frac{1}{\sqrt{1+x}+\sqrt[3]{1+x}}\,\mathrm{d}x = 2\sqrt{1+x} - 3\sqrt[3]{1+x} + 6\sqrt[6]{1+x} - 6\,\mathrm{Log}\,|1+\sqrt[6]{1+x}| + \mathrm{C}.$$

Il ne faut pas oublier, à la fin du calcul, de revenir à la variable x en remplaçant partout y par sa valeur.

Deuxième exemple

Cet exemple donne en fait une méthode générale de calcul des primitives de fractions rationnelles en *sinus* et *cosinus*. Soient P et Q deux polynômes de deux variables, à coefficients réels. On cherche à intégrer une fonction de la forme $\mathrm{P}(\cos x, \sin x)/\mathrm{Q}(\cos x, \sin x)$, et pour cela on rappelle les formules de trigonométrie suivantes : si l'on pose $t = \tan(x/2)$, alors

$$\sin x = \frac{2t}{1+t^2} \quad \text{et} \quad \cos x = \frac{1-t^2}{1+t^2}.$$

On obtient alors facilement le $\mathrm{d}x$, qui est donné par

$$\mathrm{d}x = \frac{2\,\mathrm{d}t}{1+t^2},$$

et l'on en déduit que

$$\int \frac{\mathrm{P}(\cos x, \sin x)}{\mathrm{Q}(\cos x, \sin x)}\,\mathrm{d}x = \int \frac{\mathrm{P}((1-t^2)(1+t^2)^{-1}, (2t)(1+t^2)^{-1})}{\mathrm{Q}((1-t^2)(1+t^2)^{-1}, (2t)(1+t^2)^{-1})}\,\frac{2\,\mathrm{d}t}{1+t^2},$$

qui est une intégrale de fraction rationnelle en t, et que l'on est donc en mesure, en théorie, de calculer.

À titre d'exercice, la lectrice pourra faire la même chose que ci-dessus pour une fraction rationnelle en ch et sh.

5.3. Appendice : irrationalité de e et de π

Les nombres réels sont obtenus traditionnellement par un procédé de « complétion » de l'ensemble des nombres rationnels. Ces derniers sont de la forme p/q, où p et q sont des entiers relatifs. Comme $\mathbf{Z}$ est un ensemble dénombrable, le corps $\mathbf{Q}$ des nombres rationnels est lui-même dénombrable. Il est cependant assez facile de démontrer que $\mathbf{R}$ n'est pas dénombrable, et cela sera démontré dans le chapitre sur les ensembles de mesure nulle. Ainsi $\mathbf{R}$ est-il « beaucoup plus infini » que $\mathbf{Q}$. En particulier, il existe des nombres réels non rationnels, que l'on appelle *nombres irrationnels.* Un tel argument de cardinalité ($\mathbf{Q}$ est inclus dans $\mathbf{R}$, $\mathbf{R}$ est plus gros que $\mathbf{Q}$ donc il y a des éléments de $\mathbf{R}$ qui ne sont pas dans $\mathbf{Q}$) n'est cependant pas constructif : il ne donne aucun exemple de nombre irrationnel. Il était cependant connu depuis l'antiquité grecque, cinq-cents ans avant notre ère, qu'un nombre [1] comme $\sqrt{2}$, longueur de la diagonale d'un carré de côté unité, n'est pas rationnel. La démonstration peut s'effectuer par des méthodes élémentaires si l'on raisonne par l'absurde : supposons que $\sqrt{2}=p/q$ avec p et q sans facteur commun. Alors $p^2=2q^2$ et donc 2 divise p et ainsi $4p'^2=2q^2$ et donc 2 divise q, ce qui est impossible puisque p et q sont sans facteur commun. On peut également donner de ce résultat une preuve géométrique basée sur un procédé de subdivision d'un rectangle dont les côtés sont en rapport $\sqrt{2}$.

Il a fallu attendre vingt siècles encore pour démontrer que π (périmètre d'un demi-cercle de rayon unité) est un nombre irrationnel (Johann-Heinrich Lambert, 1766).

Le nombre $\sqrt{2}$ est certes irrationnel, mais il a la propriété de vérifier (par définition) $\sqrt{2}^{\,2}-2=0$. En d'autres termes, $\sqrt{2}$ est racine d'un polynôme à coefficients entiers. On dit que $\sqrt{2}$ est un *nombre algébrique.* L'ensemble des polynômes à coefficients entiers est dénombrable, et chaque polynôme ayant un nombre fini de racines, l'ensemble des nombres réels algébriques est lui-aussi dénombrable. Par le même argument

1. Attention : pour les anciens, il n'est pas de *nombre* tel que $\sqrt{2}$. L'énoncé antique concernait les *proportions* et s'écrivait plutôt : la diagonale d'un carré n'est pas commensurable à son côté.

de cardinalité que ci-dessus, il existe des nombres réels qui ne sont pas algébriques, et que l'on appelle des *nombres transcendants.* La construction explicite de nombres transcendants est beaucoup plus tardive que celle de nombre irrationnels et est due à Joseph Liouville en 1844. Mais là encore, on aimerait connaître le caractère transcendant ou algébrique de constantes issues de la géométrie ou de l'analyse, comme π ou e, base des logarithmes naturels. C'est Charles Hermite qui le premier, en 1873, donne une telle démonstration de transcendance, pour le nombre e. S'inspirant des méthodes de Hermite, Ferdinand von Lindemann prouve en 1882 la transcendance de π. Pour illustrer la difficulté de ces questions, notons par exemple que la question de l'irrationalité de $\mathrm{e} + \pi$ ou de $\mathrm{e}\pi$ est ouverte ! Nous allons voir que les méthodes analytiques et en particulier l'emploi d'intégrales, offrent de puissants outils pour s'attaquer à ces questions, en donnant la preuve que e est un nombre transcendant.

Commençons par quelques définitions.

Définition 5.1 — *Un nombre complexe est dit algébrique s'il est racine d'un polynôme à coefficients rationnels. Il est dit transcendant sinon.*

Il se trouve que l'ensemble $\bar{\mathbf{Q}}$ des nombres algébriques forme un sous-corps du corps $\mathbf{C}$, qu'on appelle la clôture algébrique de $\mathbf{Q}$ (dans $\mathbf{C}$). Comme nous l'avons remarqué plus haut, ce corps est dénombrable.

Définition 5.2 — *Un corps* $\mathbf{K}$ *est dit algébriquement clos si tout polynôme non constant à coefficients dans* $\mathbf{K}$ *admet au moins une racine dans* $\mathbf{K}$.

Le théorème fondamental de l'algèbre (ou théorème de d'Alembert-Gauß) affirme que $\mathbf{C}$ est algébriquement clos. Ainsi tout polynôme à coefficients rationnels scinde dans $\mathbf{C}$ et donc la *clôture algébrique* du corps $\mathbf{Q}$, plus petit corps algébriquement clos contenant $\mathbf{Q}$, peut être vue comme un sous-corps du corps $\mathbf{C}$. Il se trouve que ce sous-corps est constitué exactement de l'ensemble des nombres algébriques.

Irrationalité de e

Avant de démontrer la transcendance de e, montrons, de deux manières, qu'il est irrationnel. La preuve la plus simple est de nature géométrique et est due à Jonathan Sandow en 2006. Elle se base sur l'identité suivante, que l'on doit à Leonhardt Euler :

$$\mathrm{e} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \cdots = \sum_{k=0}^{+\infty} \frac{1}{k!}.$$

On construit itérativement une suite d'intervalles $(\mathrm{I}_n)_{n\geqslant 1}$ de la manière suivante : $\mathrm{I}_1 = [2, 3]$, et l'on divise I_1 en $2 = 1 + 1$ segments égaux, et le deuxième de ces deux intervalles est $\mathrm{I}_2 = [5/2!, 6/2!]$; ensuite nous divisons I_2 en $3 = 2+1$ intervalles égaux, et $\mathrm{I}_3 = [16/3!, 17/3!]$ est le deuxième de ces intervalles. On continue ainsi, et l'on divise l'intervalle I_n en $n + 1$ segments égaux, et le deuxième de ceux-ci est I_{n+1}. Les intervalles I_n sont clairement emboîtés, et I_n est de longueur $\frac{1}{n!}$. L'intersection des I_n est

donc réduite à un unique point qui, d'après l'identité ci-dessus, n'est autre que e. Mais on a mieux : I_{n+1} est inclus dans l'intérieur de I_n, qui est de la forme $[a/n!, (a+1)/n!]$. Ainsi aucun point de I_{n+1} n'est une fraction de dénominateur $n!$. Cela implique que e ne peut, pour aucun $n \geqslant 0$, s'écrire comme une fraction avec dénominateur $n!$, ce qui entraîne que e est irrationnel (remarquer que $p/q = p(q-1)!/q!$).

Une autre preuve classique fait appel à des intégrales et permet de rentrer dans le cercles d'idées de la preuve de transcendance. La base de ces preuves est toujours le fait suivant :

Fait — il n'existe pas de suite d'entiers strictement positifs $(w_n)_{n\in\mathbf{N}}$ tendant vers 0 quand n tend vers l'infini.

Pour prouver l'irrationalité (puis la transcendance) de e nous allons construire, à partir d'intégrales de polynômes, une suite strictement décroissante de nombres positifs et, sous l'hypothèse absurde que e est rationnel, cette suite sera à valeurs entières. On pose, pour tout entier $n \geqslant 0$:

$$f_n(x) = \frac{x^n(1-x)^n}{n!}.$$

Si l'on développe f_n en monômes avec la formule du binôme, on démontre aisément les faits suivants :

- f_n est à coefficient entiers ;
- pour tout $k \geqslant 0$, $f_n^{(k)}(0)$ et $f_n^{(k)}(1)$ sont des entiers (il suffit de remarquer que 0 est une racine multiple d'ordre n de f_n) ;
- pour tout $x \in [0,1]$, on a $|f_n(x)| \leqslant 1/n!$.

Soit maintenant u un entier strictement positif et montrons par l'absurde que e^u n'est pas rationnel. Supposons donc $\mathrm{e}^u = a/b \in \mathbf{Q}$, et posons

$$\mathrm{F}_n(x) = \sum_{k=0}^{2n} (-1)^k u^{2n-k} f_n^{(k)}(x).$$

On a alors

$$\frac{\mathrm{d}}{\mathrm{d}x}[\mathrm{e}^{ux}\mathrm{F}_n(x)] = u^{2n+1}\mathrm{e}^{ux} f_n(x),$$

et donc

$$w_n = b\int_0^1 u^{2n+1}\mathrm{e}^{ux} f_n(x)\,\mathrm{d}x = a\mathrm{F}_n(1) - b\mathrm{F}_n(0)$$

est, d'après les propriétés de f_n, un nombre entier strictement positif. Mais comme $f_n(x) \leqslant 1/n!$ sur $[0,1]$, la suite (w_n) tend vers 0 quand n tend vers l'infini, ce qui impliquerait que pour n assez grand on ait $w_n < 1$, ce qui est impossible. Nous avons donc abouti à une contradiction, et démontré le théorème suivant :

Théorème 5.8 — *Pour tout nombre rationnel non nul r, e^r est irrationnel.*

En effet, si $e^{p/q}$ était rationnel, alors $e^p = (e^{p/q})^q$ le serait aussi, ce qui, d'après ce qui précède, est impossible.

Transcendance de e

La preuve de la transcendance de e est plus élaborée que celle de l'irrationalité, et fait appel à un nombre premier auxiliaire p, qui devra être suffisamment grand (on utilise ici un autre résultat arithmétique important, le caractère infini de l'ensemble des nombres premiers). Comme pour l'irrationalité, nous faisons une démonstration par l'absurde et supposons que e est algébrique, c'est-à-dire racine d'un polynôme à coefficients rationnels :

$$a_m e^m + a_{m-1}e^{m-1} + \cdots + a_0 = 0,$$

avec $a_i \in \mathbf{Q}$ et $a_m \neq 0$. Quitte à multiplier l'équation ci-dessus par le plus petit commun multiple des dénominateurs et à simplifier par une éventuelle puissance de x, on peut supposer que les coefficients a_i sont tous entiers relatifs et que $a_0 \neq 0$. Soit maintenant p un nombre premier, et posons

$$f(x) = \frac{x^{p-1}(x-1)^p(x-2)^p \cdots (x-m)^p}{(p-1)!}\ ;$$

c'est un polynôme à coefficients rationnels, de degré $p(m+1)-1$.

Si $i \geqslant 0$ et si $j \in \{0, \ldots, m\}$, alors $f^{(i)}(j)$ est un entier. En effet, par un argument de degré, on a déjà $f^{(p(m+1))} = 0$. Si $j \in \{1, \ldots, m\}$, alors j est un zéro d'ordre p de f, et donc $f(j) = f'(j) = \cdots = f^{(p-1)}(0) = 0$. Pour $i \geqslant p$, si l'on applique la règle de Leibniz, le seul terme non nul apparaît lorsque l'on dérive exactement p fois et il sort alors un $p!/(p-1)! = p$, qui est un entier divisible par p. Si $j = 0$, le même argument nous donne la nullité des $p-2$ premières dérivées, puis apparaît, en dérivant $p-1$ fois le terme $f^{(p-1)}(0) = (-1)^p \cdots (-m)^p$. Nous verrons plus bas apparaître l'expression

$$\sum_{j=0}^{m} \sum_{i=0}^{mp+p-1} a_j f^{(i)}(j),$$

et nous venons donc de voir que celle-ci est un nombre entier congru à $a_0(-1)^p \cdots (-m)^p$.

Formons maintenant le polynôme

$$F(x) = f(x) + f'(x) + \cdots + f^{(mp+p-1)}(x)$$

et calculons

$$\frac{d}{dx} e^{-x}F(x) = e^{-x}(F'(x) - F(x)) = -e^{-x}f(x).$$

Ainsi $e^{-x}F(x)$ est une primitive de $-e^{-x}f(x)$ et l'on peut donc calculer, pour tout $j \in \{0, \ldots, m\}$, l'intégrale

$$a_j e^j \int_0^j e^{-x} f(x)\, dx = a_j e^j F(0) - a_j F(j).$$

Si l'on somme sur j ces identités, on obtient

$$\sum_{j=0}^{m} a_j \mathrm{e}^j \int_0^j \mathrm{e}^{-x} f(x)\,\mathrm{d}x = \mathrm{F}(0) \sum_{j=0}^{m} a_j \mathrm{e}^j - \sum_{j=0}^{m} a_j \mathrm{F}(j).$$

Mais par hypothèse on a $a_m \mathrm{e}^m + \cdots + a_1 \mathrm{e}^1 + a_0 = 0$ et donc cette expression est finalement égale à

$$\sum_{j=0}^{m} a_j \mathrm{e}^j \int_0^j \mathrm{e}^{-x} f(x)\,\mathrm{d}x = -\sum_{j=0}^{m} \sum_{i=0}^{mp+p-1} a_j f^{(i)}(j),$$

et nous avons vu plus haut que c'était un entier congru à $a_0(-1)^p \cdots (-m)^p$; si p est assez grand (par exemple $p \geqslant \max\{m, |a_0|\}$) alors cet entier n'est pas divisible par p et donc est *non nul.*

Mais si $0 \leqslant x \leqslant m$ on a

$$|f(x)| \leqslant \frac{m^{mp+p-1}}{(p-1)!}$$

et donc

$$\left| \sum_{j=0}^{m} a_j \mathrm{e}^j \int_0^j \mathrm{e}^{-x} f(x)\,\mathrm{d}x \right| \leqslant \sum_{j=0}^{m} |a_j| \mathrm{e}^j \int_0^j \frac{m^{mp+p-1}}{(p-1)!}\,\mathrm{d}x \leqslant \sum_{j=0}^{m} j |a_j| \mathrm{e}^j \frac{m^{mp+p-1}}{(p-1)!}$$

et ce dernier terme tend vers 0 quand p tend vers $+\infty$. Cela contredit le fait que c'est un entier non nul dès que p est assez grand, et achève la démonstration du théorème suivant :

Théorème 5.9 — *Le nombre* e *est transcendant.*

En raffinant la méthode, on peut démontrer que π est lui aussi transcendant, ce qui, au passage, montre l'impossibilité de la *quadrature du cercle :* construire à la règle et au compas à partir d'un carré donné un cercle de même circonférence.

Chapitre 6
Intégrales impropres

Intégrer des fonctions définies sur des intervalles compacts s'avère très vite insuffisant, tant en mathématiques (arithmétique, analyse...) qu'en physique. On y est en effet confronté à des intégrales du type

$$\int_0^1 \frac{\mathrm{d}x}{\sqrt{x}} \quad \text{ou} \quad \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x, \quad \text{voire} \quad \int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x.$$

Or la première fonction intégrée n'est pas définie en 0, ni bornée sur $[0,1]$, tandis que les deux autres intégrale font intervenir un intervalle non borné. Nous allons voir, dans cette section, que l'on peut étendre la définition de l'intégrale de manière à prendre en compte ces deux situations. De plus, nous énoncerons un théorème dû à Heinrich Hake, qui dit que l'intégrale de Kurzweil et Henstock contient ses intégrales impropres (c.-à-d. obtenues par passage à la limite dans les bornes).

Nous ne donnerons que peu de preuves ici, et nous bornerons le plus souvent à donner les énoncés des résultats les plus importants. En fait, tout marche comme dans le cas de l'intégration des fonctions définies sur un compact, quasiment mot pour mot. Enfin, nous nous limiterons à un intervalle de la forme $[a,+\infty[$, les autres cas s'en déduisant aisément. C'est ici avant tout la pratique, au travers d'exercices, qui compte.

6.1. Intégrabilité et intégrale sur $[a,+\infty[$

Soit f une fonction définie sur l'intervalle $[a,+\infty[$. On aimerait pouvoir étendre la définition de l'intégrabilité de f (et de son intégrale, si elle est intégrable) en utilisant des sommes de Riemann et des subdivisions pointées. L'idée la plus simple consiste à *tronquer* l'intervalle $[a,+\infty[$ à droite, de manière à se ramener à des sous-intervalles compacts. Pour cela, nous allons procéder par étapes.

– Commençons par *compactifier* $[a,+\infty[$ en lui rajoutant $+\infty$, c'est-à-dire en prenant son adhérence dans la droite achevée $\overline{\mathbf{R}}$. Ainsi, nous allons travailler sur $\mathrm{I} = [a,+\infty]$.

- Définissons les subdivisions pointées :

$$\mathcal{P} = \{([a = x_0, x_1], t_1), ([x_1, x_2], t_2), \ldots, ([x_{n-1}, x_n = +\infty], t_n)\}.$$

- Il nous faut maintenant définir la somme de Riemann de f associée à $\mathcal{P}$. L'intervalle non borné $[x_{n-1}, +\infty]$ pose problème. Commençons par étendre f à $[a, +\infty]$ en posant $f(+\infty) = 0$. Ensuite, nous allons étendre la notion de finesse de manière à ce que le dernier point de marquage t_n soit toujours égal à $+\infty$. De cette manière, le dernier terme de la somme de Riemann disparaît, et celle-ci s'écrit :

$$\mathrm{S}(f, \mathcal{P}) = \sum_{k=1}^{n-1} (x_k - x_{k-1}) f(t_k).$$

- Une jauge sur I est toujours une fonction $\delta : \mathrm{I} \longrightarrow]0, +\infty[$.
- La subdivision $\mathcal{P}$ est dite δ-fine si, pour tout $k = 1, \ldots, n-1$ on a

$$[x_{k-1}, x_k] \subset [t_k - \delta(t_k)/2, t_k + \delta(t_k)/2],$$

  et, pour le dernier intervalle,

$$x_{n-1} > 1/\delta(+\infty).$$

  Ainsi, plus l'on prend $\delta(+\infty)$ petit, plus x_{n-1} doit être grand. Remarquons que, comme $[x_{n-1}, +\infty]$ est le seul intervalle de la subdivision à contenir $+\infty$, cette définition de finesse impose forcément que $t_n = +\infty$.

Munis de ces notations et conventions, il semble naturel d'adopter la

Définition 6.1 — *La fonction $f : [a, +\infty[\longrightarrow \mathbf{R}$ est intégrable si, et seulement si, il existe un nombre $\mathrm{S} \in \mathbf{R}$ tel que, pour tout $\varepsilon > 0$, il existe une jauge δ sur $[a, +\infty]$ telle que, pour toute subdivision δ-fine $\mathcal{P}$, on ait*

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}| < \varepsilon.$$

Si f est intégrable, le nombre S ci-dessus est appelé l'intégrale de f sur $[a, +\infty[$ et est noté

$$\int_{[a,+\infty[} f, \quad ou \quad \int_a^{+\infty} f(x)\,\mathrm{d}x.$$

On dit enfin que f est Lebesgue-intégrable sur $[a, +\infty[$ si f et $|f|$ sont intégrables sur $[a, +\infty[$.

L'intégrale sur $[a, +\infty[$ jouit de propriétés analogues à celles sur un intervalle compact. Bien sûr, du fait que l'intervalle est non borné (ou que la fonction est non bornée sur un intervalle borné), déterminer si la fonction est intégrable peut s'avérer problématique. Voici quelques exemples de fonctions intégrables :

$$\frac{1}{\sqrt{t}} \text{ sur } [0,1], \quad \frac{1}{x^2} \text{ sur } [1, +\infty], \quad \frac{x \sin x}{1 + x^2} \text{ sur } [0, +\infty].$$

Avant de voir comment les théorèmes vrais pour l'intégrale sur un intervalle compact se généralisent au cas étudié ici, énonçons le théorème de Hake.

Théorème 6.1 — *La fonction $f : [a, +\infty[\longrightarrow \mathbf{R}$ est intégrable si et seulement si, pour tout $c \in [a, +\infty[$, f est intégrable sur $[a, c]$ et si la limite*

$$\lim_{c \to +\infty} \int_a^c f(x)\,\mathrm{d}x = \mathrm{L}$$

existe. On a alors

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \mathrm{L}.$$

Ce théorème est admis ici, même si sa preuve, un peu délicate, est tout à fait accessible : elle est donnée en appendice. Il est très utile dans les estimations ou les calculs explicites d'intégrales et c'est lui que nous utiliserons systématiquement pour démontrer qu'une fonction est intégrable. Ainsi, combiné au théorème d'intégration par parties, il permet de démontrer que $x^{-1} \sin x$ est intégrable sur $[0, +\infty[$.

Remarque : l'intégrale de Riemann proprement dite ne permet pas de traiter les intervalles non compacts ni les fonctions non bornées, et l'intégrale de Lebesgue ne possède pas d'intégrale semi convergente comme

$$\int_0^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x.$$

6.2. Extension de quelques théorèmes aux intégrales impropres

Nous revenons ici sur quelques résultats démontrés dans le cas d'une fonction définie sur un intervalle compact et les étendons au cas non compact. Les démonstrations ne sont pas données, mais, soit elles se calquent sur celles déjà faites, soit elles sont une application facile du théorème de Hake.

La première remarque à faire est que, comme dans le cas compact, le nombre S intervenant dans la définition de l'intégrabilité est, s'il existe, unique. Le critère de Cauchy 4.1, ainsi que le théorème d'encadrement 4.2 passent sans changement. De même, l'intégrale est linéaire, positive et vérifie la relation de Chasles. Celle-ci permet de définir des intégrales entre moins l'infini et plus l'infini, par exemple :

Proposition 6.2 — *La fonction $f : \mathbf{R} \longrightarrow \mathbf{R}$ est intégrable si, et seulement si, la limite*

$$\lim_{a \to -\infty, b \to +\infty} \int_a^b f(x)\,\mathrm{d}x = \mathrm{L}$$

existe. On note dans ce cas $f \in \mathcal{I}(\mathbf{R})$ et l'on pose

$$\mathrm{L} = \int_{\mathbf{R}} f(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x.$$

Toutes sortes de variations sur ce thème sont bien sûr valables.

De même, on peut baser les intégrales indéfinies en une borne infinie de l'intervalle de définition, et le théorème 5.1 est vrai avec $a = -\infty$. La formule de la moyenne 5.5 est valable avec $b = +\infty$. Enfin, les autres résultats de la section « Intégrales et primitives » se généralisent, avec quelques modifications, au cas de bornes infinies. On peut ainsi énoncer la formule d'intégration par parties pour des fonctions définies et dérivables sur $[a, +\infty[$ de la manière suivante : soient F, G deux fonctions dérivables sur $[a, +\infty[$, et telles que $\mathrm{F}(x)\mathrm{G}(x)$ admette une limite finie quand x tend vers $+\infty$. Si FG' est intégrable sur $[a, +\infty[$, alors $\mathrm{F}'\mathrm{G}$ l'est et

$$\int_a^{+\infty} \mathrm{F}'(x)\mathrm{G}(x)\,\mathrm{d}x = \lim_{x\to+\infty} [\mathrm{FG}]_a^x - \int_a^{+\infty} \mathrm{F}(x)\mathrm{G}'(x)\,\mathrm{d}x.$$

Pour finir, notons le fait commode suivant : soit f une fonction définie sur un intervalle borné ou non $[a, b]$. On peut prolonger f à $\mathbf{R}$ tout entier en posant $\bar{f}(x) = 0$ pour tout $x \notin [a, b]$ et $\bar{f}(x) = f(x)$ pour $x \in [a, b]$.

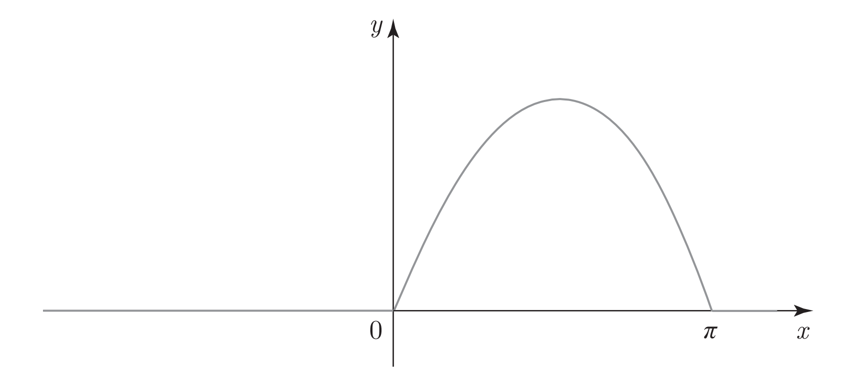


Figure 6.1 – Prolongement du sinus sur $[0, \pi]$ à $\mathbf{R}$ entier

On a alors, du fait de la relation de Chasles, le fait suivant :

Fait — La fonction f est intégrable sur $[a, b]$ si, et seulement si $\bar{f}$ est intégrable sur $\mathbf{R}$. La même chose est vraie avec l'intégrabilité au sens de Lebesgue.

6.3. Critères de convergence

Commençons par introduire la terminologie historique (et encore employée de nos jours) concernant les intégrales sur des intervalles non compacts.

Soit f une fonction définie sur l'intervalle $[a, +\infty[$. On dit que f est *localement intégrable* sur l'intervalle $[a, +\infty[$ si elle est intégrable sur tout intervalle *compact* contenu dans $[a, +\infty[$. Cela revient à dire, d'après la relation de Chasles, que pour tout $\mathrm{X} \in [a, +\infty[$, f est intégrable sur $[a, \mathrm{X}]$.

Définition 6.2 — *Soit f une fonction localement intégrable sur $[a, +\infty[$. L'intégrale*

$$\mathrm{I} = \int_a^{+\infty} f(x)\,\mathrm{d}x$$

est dite convergente *si f est intégrable sur $[a, +\infty[$. Elle est dite* absolument convergente *si f et $|f|$ sont intégrables sur $[a, +\infty[$, c'est-à-dire si f est Lebesgue-intégrable sur $[a, +\infty[$. On dit qu'elle est* semi-convergente *si elle est convergente mais pas absolument convergente. Si elle n'est pas convergente, on dit qu'elle est* divergente. *Le cas des autres types d'intervalle donne lieu à une terminologie analogue.*

Maintenant, on aimerait disposer de critères assurant la convergence d'une intégrale impropre. Il s'agit d'appliquer le théorème de comparaison 4.2, valable aussi bien dans le cas d'intervalles compacts que non compacts. Beaucoup d'exercices se posent en ces termes : « l'intégrale

$$\mathrm{I} = \int_a^{+\infty} f(x)\,\mathrm{d}x$$

est-elle *convergente ?* ». L'idée pour y répondre est d'essayer de contrôler la taille de la fonction f au voisinage de $+\infty$ et de la comparer à celle de fonctions simples. On applique alors les propositions connues sous le nom de *critères de convergence.*

Proposition 6.3 (Critère de convergence en un point infini) — *Soit f une fonction localement intégrable sur l'intervalle $[a, +\infty[$.*

1. *Si $|f(x)| \leqslant x^{-\alpha}$ pour un certain $\alpha > 1$ et pour tout x assez grand, alors* I *est absolument convergente.*
2. *Si $|f(x)| \leqslant \exp(-\eta x)$ pour un certain $\eta > 0$ et pour tout x assez grand, alors* I *est absolument convergente.*
3. *Si $|f(x)| \geqslant x^{-\alpha}$ pour un certain $\alpha \leqslant 1$ et pour tout x assez grand, alors* I *est divergente.*

On peut énoncer des critères analogues pour la convergence (non nécessairement absolue), mais il faut pour cela procéder à des encadrements de f, et non plus à une majoration de $|f|$.

Exemple : Montrons que

$$\mathrm{J} = \int_1^{+\infty} \frac{\sin x}{x}\,\mathrm{d}x$$

converge. La proposition ci-dessus ne s'applique pas directement, car la majoration brutale

$$\frac{|\sin x|}{x} \leqslant x^{-1},$$

ne rentre dans aucune case. Remarquons cependant que l'on veut intégrer un produit de deux fonctions que l'on sait aussi bien dériver qu'intégrer. Si, en particulier, on

dérive le x^{-1}, on va tomber sur un x^{-2} qui lui pourrait nous permettre d'appliquer le théorème. Fixons donc $\mathrm{X} > 1$ et intégrons par partie :

$$\int_1^{\mathrm{X}} \frac{\sin x}{x}\,\mathrm{d}x = \left[\frac{-\cos x}{x}\right]_1^{\mathrm{X}} - \int_1^{\mathrm{X}} \frac{-\cos x}{-x^2}\,\mathrm{d}x,$$

ce qui permet de conclure : le premier terme du second membre converge vers $-\cos(1)$, tandis que le second terme tend vers

$$-\int_1^{+\infty} \frac{\cos x}{x^2}\,\mathrm{d}x$$

qui existe car $|x^{-2}\cos x| \leqslant x^{-2}$. On aurait tout aussi bien pu appliquer directement le théorème d'intégration par parties énoncé un peu plus haut, mais il est tout aussi simple, et plus prudent, de l'appliquer entre des bornes finies puis d'invoquer le théorème de Hake. En guise d'exercice, vous pouvez démontrer que, si J est convergente, en revanche $|x^{-1}\sin x|$ n'est pas intégrable sur $[1, +\infty[$: l'intégrale J est semi-convergente.

On peut bien sûr énoncer la proposition ci-dessus pour le problème de la convergence en une borne finie de l'intervalle d'intégration.

Proposition 6.4 (Critères de convergence en un point fini) — *Soit f une fonction localement intégrable sur $]a, b]$.*

1. *Si $|f(x)| \leqslant 1/(x-a)^\alpha$ pour un certain $\alpha < 1$ et pour tout x assez proche de a, alors* I *est absolument convergente.*
2. *Si $|f(x)| \leqslant |\eta \operatorname{Log}|x-a||$ pour un certain $\eta > 0$ et pour tout x assez proche de a, alors* I *est absolument convergente.*
3. *Si $|f(x)| \geqslant 1/(x-a)^\alpha$ pour un certain $\alpha \geqslant 1$ et pour tout x assez proche de a, alors* I *est divergente.*

Pour finir ce chapitre, le théorème suivant permet de ramener l'étude de certaines intégrales à l'étude de séries, ce qui permet parfois de simplifier les problèmes, mais amène surtout à de jolies formules.

Théorème 6.5 (Comparaison séries-intégrales) — *Soit f une fonction* positive *sur l'intervalle $[a, +\infty[$. Pour que l'intégrale*

$$\mathrm{I} = \int_a^{+\infty} f(x)\,\mathrm{d}x$$

soit (absolument) convergente, il faut et il suffit qu'il existe une suite croissante (x_n), tendant vers $+\infty$ et telle que la série de terme général

$$u_n = \int_{x_n}^{x_{n+1}} f(x)\,\mathrm{d}x$$

soit (absolument) convergente.

6.4. Appendice : preuve du théorème de Hake

Nous avons, dans la section sur les intégrales impropres, énoncé le théorème de Hake sous la forme suivante :

Théorème 6.6 — *La fonction $f : [a, +\infty[\longrightarrow \mathbf{R}$ est intégrable si, et seulement si, pour tout $c \in [a, +\infty[$, f est intégrable sur $[a, c]$, et si la limite*

$$\lim_{c \to +\infty} \int_a^c f(x)\,\mathrm{d}x = \mathrm{L}$$

existe. On a alors

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \mathrm{L}.$$

Comme nous n'avons pas développé la théorie générale de l'intégrabilité sur les intervalles infinis outre mesure, nous allons démontrer ce théorème en remplaçant la borne $+\infty$ par un réel fini $b > a$. Quitte à poser $f(b) = 0$, nous obtenons l'énoncé suivant :

Proposition 6.7 — *Soit $f : [a, b] \longrightarrow \mathbf{R}$ une fonction numérique. Alors f est intégrable sur $[a, b]$ si, et seulement si, elle est intégrable sur tout sous-intervalle de la forme $[a, c]$, avec $a < c < b$, et si la limite*

$$\lim_{c \to b} \int_a^c f(x)\,\mathrm{d}x = \mathrm{L}$$

existe. Dans ce cas, on a

$$\int_a^b f(x)\,\mathrm{d}x = \mathrm{L}.$$

La preuve dans le cas d'une borne infinie n'est guère différente de celle que nous allons donner maintenant. Nous utiliserons en cours de route le lemme de Henstock que nous démontrerons plus loin (théorème 8.1)

Démonstration — Commençons par démontrer l'implication directe, et supposons donc f intégrable sur $[a, b]$. La relation de Chasles dit que f est intégrable sur $[a, c]$ pour tout $c \in [a, b]$, et le théorème de continuité des intégrales indéfinies permet de conclure que

$$\lim_{c \to b} \int_a^c f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x,$$

ce qui est la conclusion recherchée.

Montrons maintenant l'implication réciproque, et reprenons les notations du théorème. Soit (c_k) une suite strictement croissante telle que $a = c_0$ et $b = \lim_{k \to +\infty} c_k$. Soit également $\varepsilon > 0$. Il existe $r \in \mathbf{N}$ tel que, pour tout $k \geqslant r$, on ait

$$b - c_k \leqslant \frac{\varepsilon}{|f(b)| + 1} \quad \text{et} \quad \left| \int_a^t f - \mathrm{L} \right| \leqslant \varepsilon,\ \forall t \in [c_r, b[.$$

Pour tout $k \in \mathbf{N}$, f est intégrable sur $\mathrm{I}_k = [c_{k-1}, c_k]$, et il existe donc une jauge δ_k sur I_k telle que, pour toute subdivision δ_k-fine $\mathcal{P}_k$ de I_k, on ait

$$\left|\mathrm{S}(f, \mathcal{P}_k) - \int_{c_k}^{c_{k+1}} f\right| \leqslant \frac{\varepsilon}{2^k}.$$

Quitte à diminuer les jauges δ_k, on peut supposer que

1. $\delta_1(c_0) \leqslant \frac{1}{2}(c_1 - c_0)$;
2. $\delta_{k+1}(c_k) \leqslant \min\{\delta_k(c_k), \frac{1}{2}(c_k - c_{k-1}), \frac{1}{2}(c_{k+1} - c_k)\}$;
3. $\delta_k(t) \leqslant \min\{\frac{1}{2}(t - c_{k-1}), \frac{1}{2}(c_k - t)\}$, $t \in]c_{k-1}, c_k[$.

La famille $\{[c_{k-1}, c_k[,\ k \geqslant 1\}$ est une partition de $[a, b[$, on peut donc définir la jauge δ sur $[a, b]$ par

$$\delta(t) = \delta_k(t) \text{ si } t \in [c_{k-1}, c_k[, \text{ et } \delta(b) = b - c_r.$$

Soit $\mathcal{P}$ une subdivision δ-fine. Par construction, le point b est point de marquage pour le dernier intervalle $[x_{n-1}, b]$ de la subdivision. Par ailleurs, on a $c_r = b - \delta(b) \leqslant x_{n-1}$, car $[x_{n-1}, b] \subset [b - \delta(b)/2, b]$. Si donc s est le premier entier pour lequel $c_s \geqslant x_{n-1}$, alors $s \geqslant r$. Par hypothèse sur δ, tout intervalle de $\mathcal{P}$ contenant un c_k admet c_k pour point de marquage. Quitte à faire un coupage-collage le long des ponts de marquage, on peut supposer que les points $c_0, \dots, c_{s-1}$ sont aussi des bornes d'intervalles de $\mathcal{P}$. On peut donc restreindre $\mathcal{P}$ aux intervalles $[c_{k-1}, c_k]$, et l'on pose

$$\mathcal{Q}_1 = \mathcal{P} \cap [c_0, c_1], \dots, \mathcal{Q}_{s-1} = \mathcal{P} \cap [c_{s-2}, c_{s-1}],\ \mathcal{Q}_s = \mathcal{P} \cap [c_{s-1}, x_{n-1}],$$
$$\mathcal{Q}_{s+1} = \{([x_{n-1}, b], b)\}.$$

Chaque subdivision $\mathcal{Q}_k$, $k = 1, \dots, s-1$, est δ_k-fine, et ainsi

$$\left|\mathrm{S}(f, \mathcal{Q}_k) - \int_{c_{k-1}}^{c_k} f\right| \leqslant \frac{\varepsilon}{2^k}.$$

Enfin, $\mathcal{Q}_s$ est une *sous-subdivision* δ_s-fine de $[c_{s-1}, c_s]$, et le lemme de Henstock (*voir* le chapitre sur les théorèmes de convergence) implique que

$$\left|\mathrm{S}(f, \mathcal{Q}_s) - \int_{c_{s-1}}^{x_{n-1}} f\right| \leqslant \frac{\varepsilon}{2^s}.$$

Pour $\mathcal{Q}_{s+1}$, on a $|\mathrm{S}(f, \mathcal{Q}_{s+1})| = |f(b)|(b - x_{n-1})$, qui, par construction de r, est plus petit que ε. Comme on a

$$\mathrm{S}(f, \mathcal{P}) = \mathrm{S}(f, \mathcal{Q}_1) + \cdots + \mathrm{S}(f, \mathcal{Q}_s) + \mathrm{S}(f, \mathcal{Q}_{s+1}),$$

il vient

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{L}| \leqslant \left|\sum_{k=1}^{s} \mathrm{S}(f, \mathcal{Q}_k) - \int_a^{x_{n-1}} f\right| + \left|\mathrm{S}(f, \mathcal{Q}_{s+1})\right| + \left|\int_a^{x_{n-1}} f - \mathrm{A}\right| \leqslant 3\varepsilon.$$

Cela achève la démonstration du théorème. □

Partie II

Intégrale de Lebesgue, théorèmes de convergence et séries de Fourier

This page intentionally left blank

Chapitre 7

Ensembles de mesure nulle et notion de « presque partout »

Nous entrons maintenant dans le vif du sujet, à savoir *la théorie de l'intégration*. La première étape consiste à définir une classe d'ensembles, dits de mesure nulle, qui jouent un rôle particulier : ils sont négligeables dans les problèmes d'intégration. Ainsi, nous verrons que les ensembles dénombrables sont de mesure nulle, et que l'on peut changer les valeurs d'une fonction sur un ensemble de mesure nulle sans changer son caractère intégrable ni son intégrale. Cela confère à l'intégrale une souplesse inestimable dans les applications.

7.1. Généralités

Définition 7.1 — *Une partie* N *de* **R** *est de mesure nulle si pour tout* $\varepsilon > 0$*, il existe une suite d'intervalles ouverts* $(\mathrm{J}_k)_{k\in\mathbf{N}}$*, telle que*

$$\mathrm{N} \subset \bigcup_{k\in\mathbf{N}} \mathrm{J}_k \quad \textit{et} \quad \sum_{k=0}^{+\infty} |\mathrm{J}_k| \leqslant \varepsilon.$$

Une propriété $\mathrm{P}(x)$ *dépendant de* x *pour* x *dans une partie* A *de* **R** *est dite vraie* presque partout *(ou* pour presque tout $x \in \mathrm{A}$*) si l'ensemble des* x *de* A *tels que* $\mathrm{P}(x)$ *soit violée est une partie de mesure nulle.*

On dit par exemple que $f : [a, b] \longrightarrow \mathbf{R}$ est nulle presque partout si $\{x \in [a, b],\ f(x) \neq 0\}$ est de mesure nulle.

Résumons quelques propriétés fondamentales des ensembles de mesure nulle.

Proposition 7.1 — *Une partie dénombrable est de mesure nulle. Une union dénombrable d'ensembles de mesure nulle est de mesure nulle. Une partie incluse dans un ensemble de mesure nulle est de mesure nulle. Un intervalle d'intérieur non vide n'est pas de mesure nulle.*

Nous laissons la démonstration en exercice pour le lecteur. Le fait qu'un intervalle non trivial n'est pas de mesure nulle est assez délicat à démontrer directement, mais c'est par contre un corollaire facile du théorème 7.2 qui est démontré ci-dessous.

7.2. Intégrale et ensembles de mesure nulle

Avec un peu plus de travail, nous pouvons démontrer le théorème suivant, qui relie la notion d'intégrale à celle de mesure nulle.

Théorème 7.2 — *Soit f une fonction nulle presque partout sur $[a,b]$. Alors f est intégrable, d'intégrale nulle.*

DÉMONSTRATION — Notons E l'ensemble de mesure nulle où f est non nulle. L'ensemble E est l'union dénombrable des ensembles de mesure nulle $\mathrm{E}_m = \{x, m-1 < |f(x)| \leqslant m\}$, et donc, étant donné $\varepsilon >$ et $m \geqslant 0$, il existe une collection d'intervalles *ouverts* $\{\mathrm{J}_k^m, k \in \mathbf{N}\}$, tels que

$$\mathrm{E}_m \subset \bigcup_k \mathrm{J}_k^m \text{ et } \sum_{k=0}^{+\infty} |\mathrm{J}_k^m| \leqslant \frac{\varepsilon}{m2^m}.$$

Si $x \in \mathrm{E}$, il existe un unique entier $m(x)$ tel-que $x \in \mathrm{E}_m$, et un entier minimal $k(x)$ tel que $x \in \mathrm{J}_k^m(x)$. On définit alors une jauge δ sur I par

$$\delta(x) = 1 \text{ si } x \notin \mathrm{E}, \;\; \delta(x) = \frac{1}{2}\mathrm{dist}\left(x, \partial \mathrm{J}_{k(x)}^{m(x)}\right) \text{ si } x \in \mathrm{E},$$

de telle sorte que si $x \in \mathrm{E}$, l'intervalle $[x-\delta(x), x+\delta(x)]$ soit inclus dans J_k^m (la notation $\mathrm{dist}(x,\partial\mathrm{J})$ désigne la distance entre x et le bord de $\mathrm{J} = [c,d]$, c.-à-d. $\min\{|x-c|, |x-d|\}$). Soit maintenant $\mathcal{P} = \{(\mathrm{I}_l, t_l)\}$ une subdivision pointée δ-fine. Alors

$$\mathrm{S}(f,\mathcal{P}) = \sum_{l=1}^{n} f(t_l)|\mathrm{I}_l| = \sum_{t_l \in \mathrm{E}} f(t_l)|\mathrm{I}_l| \leqslant \sum_m \sum_k \sum_{t_l \in \mathrm{Z_E} \cap \mathrm{J}_k^m} f(t_l)|\mathrm{I}_l|.$$

Mais, pour les l tels que $t_l \in \mathrm{E}_m \cap \mathrm{J}_k^m$, les I_l sont disjoints, de longueur totale plus petite que $|\mathrm{J}_k^m|$, et $|f(t_l)| < m$, donc on en déduit que

$$|\mathrm{S}(f,\mathcal{P})| \leqslant \sum_m \sum_k m|\mathrm{J}_k^m| \leqslant \sum_m m\frac{\varepsilon}{m2^m} \leqslant \varepsilon,$$

et cela achève la démonstration du théorème. □

On utilise souvent le théorème 7.2 sous la forme suivante :

Corollaire 7.3 — *Si f est intégrable sur $[a, b]$ et si g est une fonction sur $[a, b]$ égale à f presque partout, alors g est intégrable et*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x.$$

C'est une conséquence immédiate de la linéarité de l'intégrale.

On peut maintenant se poser la question d'une réciproque au théorème 7.2 : si une fonction est intégrable d'intégrale nulle sur un intervalle, doit-elle être nulle sur icelui ? Bien sûr l'exemple d'une fonction impaire (par exemple le sinus) intégrée sur un intervalle symétrique par rapport à zéro nous indique qu'un tel résultat doit éviter les compensations possibles en différents endroits de l'intervalle d'intégration. Nous devons donc nous limiter aux fonctions de signe constant. Enfin le théorème 7.2 nous indique également que ce n'est pas l'annulation partout mais seulement presque partout que l'on peut espérer :

Théorème 7.4 — *Soient $a < b$ et f une fonction définie sur $[a, b]$, intégrable, positive et d'intégrale nulle. Alors f est nulle presque partout.*

Nous admettrons ce théorème ici, car sa démonstration fait appel à un lemme de recouvrement technique, et reportons celle-ci à l'appendice. Notons cependant le cas, important, des fonctions continues :

Proposition 7.5 — *Soient $a < b$ et $f : [a, b] \longrightarrow \mathbf{R}$ une fonction continue et positive, d'intégrale nulle sur $[a, b]$. Alors f est nulle partout.*

Démonstration — Nous pourrions bien sûr invoquer le théorème 7.4, mais nous pouvons donner une preuve directe. Nous allons pour cela démontrer la contraposée du résultat : si f est continue, positive et non nulle en au moins un point x_0, alors son intégrale sur $[a, b]$ est non nulle. En effet, f étant continue en x_0, il existe $\eta > 0$ tel que pour tout $x \in [a, b] \cap \,]x_0 - \eta, x_0 + \eta[$ on ait $f(x) \geqslant f(x_0)/2 > 0$. La positivité de l'intégrale et le fait que $f \geqslant 0$ sur tout $[a, b]$ donne alors

$$\int_a^b f(x)\,\mathrm{d}x \geqslant \int_{[a,b]\cap]x_0-\eta,x_0+\eta[} f(x)\,\mathrm{d}x \geqslant \frac{f(x_0)}{2} \Big|[a, b] \cap \,]x_0 - \eta, x_0 + \eta[\Big| > 0,$$

ce qui est la conclusion recherchée. □

7.3. L'ensemble triadique de Cantor et l'escalier du diable

Nous allons ici construire un ensemble de mesure nulle qui n'est pas dénombrable, par un procédé itératif de retrait de sous-intervalles de $[0, 1]$. Cet ensemble admet une multitude de caractérisations. Lui et ses congénères, appelés génériquement *ensembles de Cantor*, sont d'une importance cruciale dans de nombreux domaines. Son inventeur, Georg Cantor, est un des fondateurs des mathématiques modernes.

La construction la plus simple de l'ensemble de Cantor K consiste à commencer par couper l'intervalle $[0, 1]$ en trois intervalles de longueur $1/3$:

$$\mathrm{I}_0^1 = [0, 1/3], \;\; \mathrm{I}_1^1 =]1/3, 2/3[, \;\; \mathrm{I}_2^1 = [2/3, 1].$$

On retire alors l'intervalle du milieu (d'où le nom anglo-saxon de *middle-third Cantor set*), et l'on pose $\mathrm{K}^1 = \mathrm{I}_0^1 \cup \mathrm{I}_2^1$. On répète ce processus de découpage et de retrait aux deux intervalles restants, I_0^1 et I_2^1. Ainsi,

$$\mathrm{I}_0^1 = \mathrm{I}_{00}^1 \cup \mathrm{I}_{01}^2 \cup \mathrm{I}_{02}^2 = [0, 1/3^2] \cup]1/3^2, 2/3^2[\cup [2/3^2, 1/3],$$

et

$$\mathrm{I}_2^1 = \mathrm{I}_{20}^2 \cup \mathrm{I}_{21}^2 \cup \mathrm{I}_{22}^2 = [2/3, 7/3^2] \cup]7/3^2, 8/3^2[\cup [8/3^2, 1].$$

On pose alors

$$\mathrm{K}^2 = \mathrm{I}_{00}^2 \cup \mathrm{I}_{02}^2 \cup \mathrm{I}_{20}^2 \cup \mathrm{I}_{22}^2.$$

L'ensemble K^3 est obtenu en retirant à chacun des 2^2 intervalles constituant K^2 l'intervalle tiers milieu. Ainsi, K^3 est l'union de 2^3 intervalles fermés disjoints, tous de longueur 3^{-3}. L'itération de ce procédé nous donne, à chaque étape, un compact K^n qui est l'union de 2^n intervalles fermés disjoints, tous de longueur 3^{-n} (*voir* la figure 7.1).

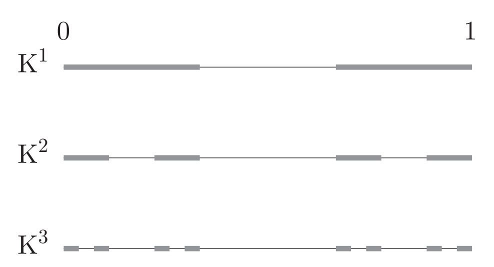


Figure 7.1 – Les trois premières étapes de construction de K

La suite des sous-ensembles K^n est décroissante, et l'ensemble de Cantor est leur intersection :

$$\mathrm{K} = \bigcap_{n \geqslant 1} \mathrm{K}^n.$$

Remarque : on peut donner une autre description de K, en termes de développement d'un nombre réel en base 3. En effet, tout nombre réel entre 0 et 1 peut s'écrire en base 3, de même qu'il admet un développement décimal. Ainsi,

$$0 = 0{,}0000\ldots,\ \ 1/3 = 0{,}1000\ldots,\ \ 2/3 = 0{,}2000\ldots,\ \ 1/2 = 0{,}1111\ldots, \text{ etc.}$$

Tout nombre $x \in [0,1]$ admet ainsi une écriture du type

$$x = \sum_{i=1}^{+\infty} \alpha_i 3^{-i},$$

où les « décimales » triadiques α_i valent 0, 1 ou 2. C'est ce développement que nous avons utilisé dans la notation des intervalles I_0^1 et autres. En effet, I_0^1 est constitué des nombres dont la première décimale triadique est un 0, I_1^1 ceux pour lesquelles celle-ci est un 1, et I_2^1 ceux qui commencent par un 2. Du coup, on code de la même manière les intervalles $\mathrm{I}^n_{\alpha_1,\alpha_2,\ldots,\alpha_n}$: c'est l'ensemble (à au plus deux exceptions près, mais passons sur ce détail) des nombres $x \in [0,1]$ tels que le développement de x en base 3 commence par $0,\alpha_1\ldots\alpha_n$. Il est alors facile de décrire K, au vu de ces remarques : c'est l'ensemble des nombres réels $x \in [0,1]$ dont le développement en base 3 *ne comporte aucun* 1.

L'ensemble K est un fermé de $[0,1]$, car c'est une intersection de fermés. Il est non vide, car il s'agit d'une intersection décroissante de fermés dans un compact. Mais on a mieux :

Théorème 7.6 — *L'ensemble de Cantor* K *est un compact de mesure nulle et non dénombrable.*

Démonstration — La suite K^n est décroissante, et chaque K^n est l'union de 2^n intervalles I_k^n, fermés et de longueur 3^{-n}. Chaque I_k^n est inclus dans un intervalle ouvert J_k^n de longueur $(2{,}9)^{-n}$, et ainsi l'on a, pour tout n,

$$\mathrm{K} \subset \mathrm{K}^n \subset \bigcup_{k=1}^{2^n} \mathrm{J}_k^n,$$

avec

$$\sum_{k=1}^{2^n} |\mathrm{J}_k^n| \leqslant \sum_{k=1}^{2^n} (2{,}9)^{-n} = \left(\frac{2}{2{,}9}\right)^n,$$

et donc K est de mesure nulle.

Montrons maintenant que K est non dénombrable. Pour cela, raisonnons par l'absurde et supposons K dénombrable. On peut alors l'énumérer : $\mathrm{K} = \{x_1, x_2, \ldots, x_n, \ldots\}$. Construisons alors un point x qui sera dans K mais pas dans l'énumération $x_1, \ldots$ Pour cela, procédons par induction, et utilisons le fait que dans la construction de K, le nombre de choix possibles dans les intervalles I_k^n croît exponentiellement vite

avec n. Tout d'abord, x_1 est soit dans I_0^1, soit dans I_2^1. Posons J_1 égal à celui des deux où x_1 *n'est pas*. Supposons maintenant construit J_{n-1}, et posons J_n égal à l'intervalle de $\mathrm{J}_{n-1} \cap \mathrm{K}^n$ où x_n *n'est pas*. On obtient ainsi une suite décroissante d'intervalles fermés non vides, dont l'intersection est donc non vide et contient un point $x \in \mathrm{K}$. Pour tout n, on a $x_n \notin \mathrm{J}_n$ et $x \in \mathrm{J}_n$, donc $x \neq x_n$ et l'on a achevé la démonstration du théorème. □

Remarquons au passage qu'il découle de ce théorème que $\mathbf{R}$ soit même n'est pas dénombrable. Une fois construit l'ensemble de Cantor, nous pouvons introduire la fonction dite « escalier du diable », qui est effectivement diabolique puisque, comme nous le verrons, même s'il permet de monter d'un étage il y a toujours, entre deux marches différentes, une infinité de marches. Construisons cette fonction par approximation. Définissons φ_1 comme étant constante égale à 2^{-1} sur I_1^1, $\varphi_1(0) = 0$ et $\varphi_1(1) = 1$, et enfin, sur les 2 intervalles restants, φ_1 est affine, de manière à ce qu'elle soit continue. Il n'y a qu'une seule fonction φ_1 vérifiant ces hypothèses (*voir* la figure 7.2 pour comprendre l'évidence de cette affirmation). De manière générale, φ_n est l'unique fonction continue, affine par morceaux, telle que $\varphi_n(0) = 0, \varphi_n(1) = 1$ et φ_n prend successivement, sur les $2^n - 1$ intervalles retirés de $[0, 1]$ pour construire K^n, les valeurs $1/2^n, 2/2^n, \ldots, (2^n - 1)/2^n$.

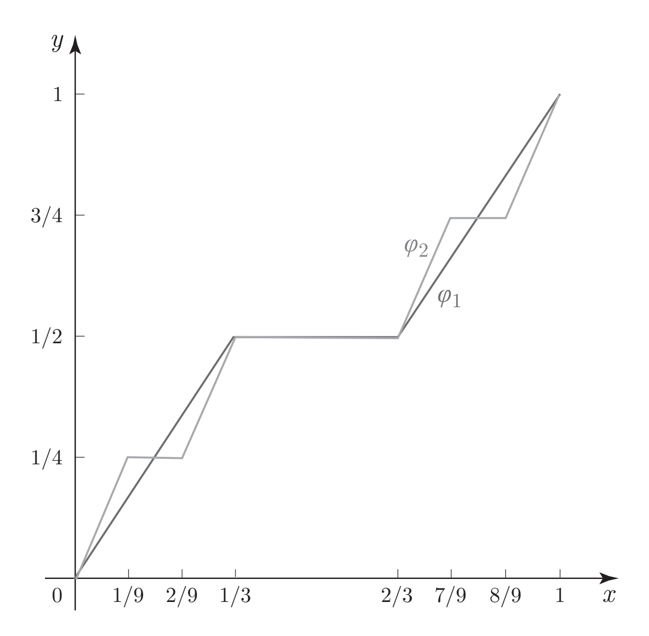


Figure 7.2 – L'escalier du diable de Cantor-Lebesgue

La suite K^n est décroissante, donc la suite $[0, 1] - \mathrm{K}^n$ est croissante, et pour tout $x \in [0, 1] - \mathrm{K}$, $\varphi_n(x)$ est constante à partir d'un certain rang $n(x)$. En fait, pour tout

$n \geqslant 1$, les graphes de φ_n et φ_{n+1} soient coïncident, soit sont dans la même bande horizontale de hauteur $1/2^n$. On en déduit que

$$\text{pour tout } x \in [0,1], \text{ pour tout } n \geqslant 1, \ |\varphi_n(x) - \varphi_{n+1}(x)| \leqslant 2^{-n}.$$

La suite $\varphi_n(x)$ est de Cauchy, et elle admet donc une limite $\varphi(x)$, qui satisfait de plus :

$$\text{pour tout } x \in [0,1], \text{ pour tout } n \geqslant 1, \ |\varphi_n(x) - \varphi(x)| \leqslant 2^{-(n-1)}.$$

Ainsi, nous avons montré que (φ_n) converge uniformément vers φ sur $[0,1]$. Comme les φ_n sont continues et croissantes, φ l'est aussi, et de plus $\varphi(0) = 0$, $\varphi(1) = 1$.

Soit maintenant x choisi hors de l'ensemble de mesure nulle K. Comme φ_n est égale à une constante indépendante de $n \geqslant n(x)$ au voisinage de x, φ est elle aussi constante au voisinage de x. En particulier, φ est dérivable en x et $\varphi'(x) = 0$.

De toute cette discussion, nous déduisons le théorème suivant :

Théorème 7.7 — *La fonction φ, dite « escalier du Diable », est continue surjective de $[0,1]$ sur $[0,1]$, croissante. Elle est par ailleurs presque partout dérivable, avec $\varphi'(x) = 0$ presque partout.*

Nous voyons ici apparaître une pathologie dans les liens entre intégration et primitivation. Ainsi, φ est presque partout dérivable, de dérivée nulle. En d'autres termes, φ est une « presque » primitive de 0. Mais si l'on appliquait le théorème fondamental à φ, on obtiendrait

$$\int_0^1 \varphi'(x)\,\mathrm{d}x = \varphi(1) - \varphi(0) = 1,$$

mais le terme de gauche est nul car φ' est nulle presque partout. Cela est impossible. On ne peut donc espérer généraliser le théorème fondamental avec des presque partout dedans. Si l'on veut faire cela, il faut faire une hypothèse supplémentaire sur les variations locales de φ.

7.4. Appendice : lemme de Vitali, application à la mesure nulle

Les théorèmes de recouvrement sont centraux dans de nombreux domaines des mathématiques, car ils permettent en quelque sorte de travailler sur une version grossière de l'espace. Ils permettent par exemple d'appliquer les idées de la géométrie différentielle à des espaces *a priori* étrangers à ce domaine. Nous allons ici démontrer l'un de ces lemmes de recouvrement, le *lemme de Vitali*, qui est d'essence géométrique : il permet de contrôler les parties bornées de $\mathbf{R}^d$ en les enserrant dans des unions disjointes de pavés ou de boules, à un ensemble « petit » près.

7.4.1. Lemme de Vitali

Nous nous contentons pour l'instant d'énoncer et démontrer le lemme en dimension $d = 1$. Nous indiquerons plus loin comment il peut se généraliser en dimension supérieure.

Définition 7.2 — *On dit qu'une famille* $\mathcal{F}$ *d'intervalles fermés* d'intérieur non vide *de* $\mathbf{R}$ *est un recouvrement de Vitali d'une partie* E *de* $\mathbf{R}$ *si pour tout* $\varepsilon > 0$ *et tout* $x \in \mathrm{E}$*, il existe* $\mathrm{I} \in \mathcal{F}$ *tel que* $x \in \mathrm{I}$ *et* $|\mathrm{I}| < \varepsilon$*.*

En d'autres termes, $\mathcal{F}$ recouvre E et tout élément de E est recouvert par des éléments de $\mathcal{F}$ de longueur arbitrairement petite ; en particulier, un recouvrement de Vitali n'est jamais fini.

Exemple 1 : soient $\mathrm{E} = [0, 1]$ et $\mathcal{F} = \{[a, b],\ a < b \text{ et } a, b \in \mathbf{Q}\}$. Cela définit un recouvrement de Vitali dénombrable de $[0, 1]$.

Exemple 2 : soient $\mathrm{E} = \{x_0\}$ et $\mathcal{F} = \{[x_0 - \varepsilon, x_0 + \varepsilon],\ \varepsilon > 0\}$.

Théorème 7.8 (Lemme de Vitali) — *Soient* E *une partie bornée de* $\mathbf{R}$ *et* $\mathcal{F}$ *un recouvrement de Vitali de* E*. Alors pour tout* $\varepsilon > 0$ *il existe une famille finie* $\{\mathrm{I}_1, \ldots, \mathrm{I}_p\}$ *d'éléments deux à deux disjoints de* $\mathcal{F}$*, ainsi qu'une famille au plus dénombrable* $(\mathrm{J}_l)_{l \geqslant p+1}$ *d'intervalles telles que :*

– E *est inclus dans l'union des* I_k *et des* J_l *:*

$$\mathrm{E} \subset \bigcup_{k=1}^{p} \mathrm{I}_k \ \cup \bigcup_{l=p+1}^{+\infty} \mathrm{J}_l\,;$$

– *les* J_l *vérifient :*

$$\sum_{l=p+1}^{+\infty} |\mathrm{J}_l| < \varepsilon.$$

En d'autres termes, à un ensemble petit près (c.-à-d. une union d'intervalles dont la somme des longueurs est arbitrairement petite) tout borné de $\mathbf{R}$ muni d'un recouvrement de Vitali peut s'écrire comme union d'une famille finie d'éléments deux à deux disjoints de ce recouvrement.

Démonstration — Nous allons construire par récurrence une suite $(\mathrm{I}_k)_{k \in \mathbf{N}}$, éventuellement finie, d'éléments de $\mathcal{F}$, qui recouvriront suffisamment E pour pouvoir conclure.

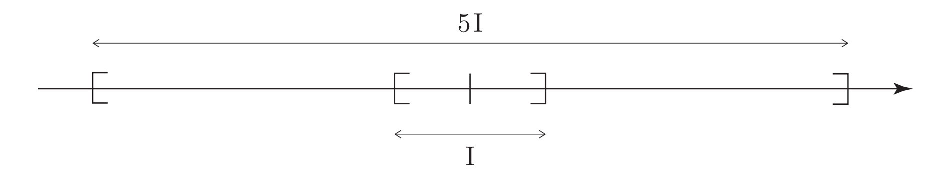
Commençons par remarquer qu'on peut toujours supposer que $\mathrm{E} \subset [0, 1]$ et que pour tout $\mathrm{I} \in \mathcal{F}$, on a $\mathrm{I} \subset [-1/2, 3/2]$, de sorte que $|\mathrm{I}| \leqslant 2$. Effectuons les premiers pas de la construction de la suite (I_k).

– Posons $\mathcal{F}_0 = \mathcal{F}$ et $L_0 = \sup\{|I|,\ I \in \mathcal{F}\} \in]0, 2]$. Soit alors $I_1 \in \mathcal{F}_0$ tel que $|I_1| \geqslant L_0/2$; si $E \subset I_1$, nous pouvons nous arrêter là, le théorème est démontré (avec $p = 1$ et J_l n'importe quelle suite d'intervalles dont la somme des longueurs est plus petite que ε.

– Sinon, soient $\mathcal{F}_1 = \{I \in \mathcal{F}_0,\ I \cap I_1 = \emptyset\}$ et $L_1 = \sup\{|I|,\ I \in \mathcal{F}_1\}$. Prenons alors $I_2 \in \mathcal{F}_2$ satisfaisant $|I_2| \geqslant L_1/2$; si $E \subset I_1 \cup I_2$, on peut s'arrêter là.

Sinon, soit $k \geqslant 1$ et supposons que nous ayons construit $I_1 \in \mathcal{F}_0, \dots, I_k \in \mathcal{F}_{k-1}$; si $E \subset I_1 \cup \cdots \cup I_k$, le procédé s'arrête et le théorème est démontré. Sinon, on pose $\mathcal{F}_k = \{I \in \mathcal{F}_{k-1},\ I \cap (I_1 \cup \cdots \cup I_k) = \emptyset\}$ et $L_k = \sup\{|I|,\ I \in \mathcal{F}_k\}$. Soit alors $I_{k+1} \in \mathcal{F}_k$ tel que $|I_{k+1}| \geqslant L_k/2$, et l'on s'arrête si E est inclus dans l'union des I_k. Nous supposerons donc désormais que le processus ci-dessus se continue indéfiniment.

Par convention, si I est un intervalle borné de $\mathbf{R}$, nous noterons 5I l'intervalle de même nature que I, qui a le même centre, mais une longueur cinq fois supérieure.



Fixons maintenant $\varepsilon > 0$; comme les I_k sont disjoints deux à deux et tous inclus dans $[-1/2, 3/2]$, la série $\sum |I_k|$ est convergente (et sa somme est plus petite que 2). il existe donc $p > 0$ tel que

$$\sum_{k=p+1}^{+\infty} |I_k| < \varepsilon/5.$$

Comme $L_k \leqslant 2|I_{k-1}|$, cela implique en particulier que $\lim_{k\to+\infty} L_k = 0$.

Soit maintenant

$$x \in E - \bigcup_{k=1}^{p} I_k.$$

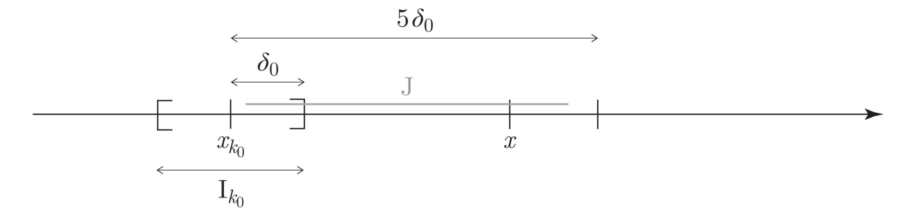
Comme les I_k sont fermés et que $\mathcal{F}$ est un recouvrement de Vitali, il existe $J \in \mathcal{F}$ tel que

$$x \in J \ \text{ et } \ J \cap \bigcup_{k=1}^{p} I_k = \emptyset\,;$$

en particulier, $J \in \mathcal{F}_p$. Nous affirmons cependant qu'il existe $k > p$ tel que $J \cap I_k \neq \emptyset$. En effet, si ce n'était pas le cas, on aurait, pour tout k, $J \in \mathcal{F}_k$ et donc $|J| \leqslant L_k$; cela entraînerait que $|J| = 0$, ce qui est impossible puisque J, étant élément de $\mathcal{F}$, est d'intérieur non vide. Nous pouvons donc considérer k_0 le plus petit entier $k > p$ tel que $J \cap I_k \neq \emptyset$; par définition, $J \in \mathcal{F}_{k_0-1}$, et J vérifie

$$|J| \leqslant L_{k_0-1} \leqslant 2|I_{k_0}|,\ J \cap I_{k_0} = \emptyset,\ x \in J.$$

Le schéma suivant :



où x_{k_0} est le milieu de I_{k_0} et $\delta_0 = |\mathrm{I}_{k_0}|/2$, montre que $x \in 5\mathrm{I}_{k_0}$.

Nous pouvons en conclure, comme x a été pris arbitrairement dans E privé de $\mathrm{I}_1 \cup \cdots \cup \mathrm{I}_p$, que, si l'on pose $\mathrm{J}_l = 5\mathrm{I}_l$ pour $l \geqslant p+1$, on a

$$\mathrm{E} \subset \bigcup_{k=1}^{p} \mathrm{I}_k \ \cup \ \bigcup_{l=p+1}^{+\infty} \mathrm{J}_l$$

ainsi que

$$\sum_{l=p+1}^{+\infty} |\mathrm{J}_l| = 5 \sum_{k=p+1}^{+\infty} |\mathrm{I}_k| < \varepsilon,$$

et le théorème est donc démontré. □

Dans la démonstration que nous venons d'achever, nous avons en fait montré que

$$\mathrm{E} - \bigcup_{k \geqslant 1} \mathrm{I}_k$$

est un ensemble de mesure nulle. Nous pouvons donc énoncer un corollaire simple au lemme de recouvrement de Vitali :

Corollaire 7.9 — *Avec les mêmes hypothèses et notations que plus haut, il existe une suite finie ou dénombrable* $(\mathrm{I}_k)_{k \geqslant 1}$ *d'éléments deux à deux disjoints de* $\mathcal{F}$ *ainsi qu'un ensemble de mesure nulle* N *tels que*

$$\mathrm{E} \subset \mathrm{N} \cup \bigcup_{k \geqslant 1} \mathrm{I}_k,$$

Remarque : le lemme de Vitali est vrai aussi en dimension plus grande, dans $\mathbf{R}^d$. Il faut cependant prendre garde à une généralisation trop naïve de la notion de recouvrement de Vitali. Le théorème est vrai pour les recouvrements constitués de boules (pour n'importe quelle norme sur $\mathbf{R}^d$), et la démonstration est alors exactement la même que celle donnée ci-dessus. Si l'on veut un énoncé faisant appel à des recouvrements par des pavés, il faut faire attention à ce que les pavés ne s'écrasent pas trop le long de l'une de leur direction de base. Pour cela, on impose, dans la définition

de recouvrement de Vitali par des pavés, que le rapport entre les longueurs des côtés des pavés reste dans un intervalle compact de $\mathbf{R}_+^*$: les pavés ressemblent alors à des cubes, distordus de manière contrôlée. Il existe enfin de vastes généralisations de ce lemme, et l'ingrédient principal est toujours un lien entre le volume (ou la mesure) d'une boule et celui de la boule de rayon double.

7.4.2. Mesure nulle et intégrale

Nous allons ici démontrer le théorème énoncé plus haut :

Théorème 7.10 — *Soit $f : [a, b] \longrightarrow \mathbf{R}$ une fonction positive et intégrable sur $[a, b]$, d'intégrale nulle. Alors f est nulle presque partout.*

Démonstration — Nous allons nous appuyer sur un lemme qui traduit que la mesure de Lebesgue est une mesure régulière et qui a pour corollaire le fait que le théorème est vrai pour les fonctions indicatrices de parties.

Lemme 7.11 — *Soit $\mathrm{S} \subset [a, b]$ une partie telle que la fonction indicatrice $\mathbf{1}_\mathrm{S}$ de S soit Lebesgue-intégrable. Alors pour tout $\varepsilon > 0$ il existe une famille dénombrable d'intervalles J_k telle que*

$$(\star) \qquad \mathrm{S} \subset \bigcup_{k=1}^{\infty} \mathrm{J}_k$$

et

$$(\dagger) \qquad \sum_{k=1}^{+\infty} |\mathrm{J}_k| \leqslant \int_a^b \mathbf{1}_\mathrm{S}(x)\,\mathrm{d}x + \varepsilon.$$

Avant de démontrer le lemme, remarquons que la conclusion implique trivialement que

$$\int_a^b \mathbf{1}_\mathrm{S}(x)\,\mathrm{d}x \leqslant \sum_{k=1}^{+\infty} |\mathrm{J}_k|$$

car, d'après $(\star)$, on a

$$\mathbf{1}_\mathrm{S}(x) \leqslant \mathbf{1}_{\cup \mathrm{J}_k}(x)$$

pour tout x. Passons maintenant à la démonstration du lemme ; notons A l'intégrale de $\mathbf{1}_\mathrm{S}$ sur $[a, b]$ et soit $\eta > 0$. Il existe une jauge δ sur $[a, b]$ telle que pour toute subdivision δ-fine $\mathcal{P}$ on ait $0 \leqslant \mathrm{S}(\mathbf{1}_\mathrm{S}, \mathcal{P}) \leqslant \mathrm{A} + \eta$. La famille d'intervalles $\mathcal{F} = \{]x - s, x + s[, s < \delta(x)/2, x \in \mathrm{S}\}$ est un recouvrement de Vitali de S, auquel on peut donc appliquer le lemme du même nom. Il existe donc un ensemble de mesure nulle $\mathcal{N} \subset \mathrm{S}$ et une famille dénombrable $]x_i - s_i, x_i + s_i[$ d'intervalles éléments de $\mathcal{F}$, deux à deux disjoints et tels que

$$\mathrm{S}' \subset \bigcup_{j=1}^{\infty}]x_i - s_i, x_i + s_i[,$$

où $\mathrm{S}' = \mathrm{S} - \mathcal{N}$. Comme $\mathbf{1}_{\mathrm{S}'} = \mathbf{1}_\mathrm{S}$ presque partout, la fonction $\mathbf{1}_{\mathrm{S}'}$ est Lebesgue intégrable sur $[a, b]$ d'intégrale A.

Soit $\mathrm{N} \geqslant 1$ un entier arbitrairement grand, et considérons la famille $\{([x_1 - s_1, x_1 + s_1], x_1), \dots, ([x_\mathrm{N} - s_\mathrm{N}, x_\mathrm{N} + s_n], x_\mathrm{N})\}$. Par construction de $\mathcal{F}$ et du fait que les intervalles ouverts la constituant sont deux à deux disjoints, cette famille est une sous-subdivision δ-fine de $[a, b]$, que l'on peut compléter en un subdivision δ-fine $\mathcal{P}$ de $[a, b]$. Comme $\mathbf{1}_\mathrm{S}$ est nulle hors de l'union des $[x_i - s_i, x_i + s_i]$, on en déduit que

$$\sum_{i=1}^{\mathrm{N}} \big|[x_i - s_i, x_i + s_i]\big| \leqslant \mathrm{S}(\mathbf{1}_\mathrm{S}, \mathcal{P}) \leqslant \mathrm{A} + \eta$$

et, comme N a été choisi arbitrairement grand, en posant $\mathrm{J}_i =]x_i - s_i, x_i + s_i[$:

$$\sum_{i=1}^{+\infty} |\mathrm{J}_i| \leqslant \mathrm{A} + \eta.$$

Ceci achève la démonstration du lemme.

Une conséquence immédiate du lemme est que si S est une partie de $[a, b]$ telle que $\mathbf{1}_\mathrm{S}$ soit intégrable d'intégrale nulle, alors S est de mesure nulle.

Nous pouvons maintenant démontrer le théorème. Posons $\mathrm{S} = \{x,\ f(x) \neq 0\}$ et, pour tout $m \in \mathbf{Z}$, définissons

$$\mathrm{S}_m = \{x,\ 2^m < f(x) \leqslant 2^{m+1}\}.$$

La famille des S_m forme une partition de S.

Soit $m \geqslant 0$ et fixons $\varepsilon > 0$. Soit δ une jauge sur $[a, b]$ telle que pour toute subdivision δ-fine on ait $0 \leqslant \mathrm{S}(f, \mathcal{P}) \leqslant \varepsilon$. Soit enfin $\mathcal{P}$ une telle subdivision. Comme $\mathbf{1}_{\mathrm{S}_m}$ est majorée par f et comme $f \geqslant 0$, on a

$$0 \leqslant \mathrm{S}(\mathbf{1}_{\mathrm{S}_m}, \mathcal{P}) \leqslant \mathrm{S}(f, \mathcal{P}) \leqslant \varepsilon,$$

ce qui prouve que $\mathbf{1}_{\mathrm{S}_m}$ est intégrable d'intégrale nulle. D'après le lemme qui précède, S_m est de mesure nulle.

Soit maintenant $m < 0$ et fixons $\varepsilon > 0$. Soit δ une jauge sur $[a, b]$ telle que pour toute subdivision δ-fine on ait $0 \leqslant \mathrm{S}(f, \mathcal{P}) \leqslant 2^m \varepsilon$. Soit enfin $\mathcal{P}$ une telle subdivision. Si $x \in \mathrm{S}_m$ on a

$$1 = \mathbf{1}_{\mathrm{S}_m}(x) \leqslant \frac{f(x)}{2^m}$$

et si $x \notin \mathrm{S}_m$ on a

$$0 = \mathbf{1}_{\mathrm{S}_{m(x)}} \leqslant \frac{f(x)}{2^m}.$$

On en déduit donc que $\mathbf{1}_{\mathrm{S}_m}$ est majorée par $2^{-m} f$ et, comme ci-dessus, on en tire

$$0 \leqslant \mathrm{S}(\mathbf{1}_{\mathrm{S}_m}, \mathcal{P}) \leqslant \mathrm{S}(2^{-m} f, \mathcal{P}) = 2^{-m} \mathrm{S}(f, \mathcal{P}) \leqslant \varepsilon$$

et donc S_m est de mesure nulle.

Nous avons donc montré que pour tout $m \in \mathbf{Z}$, S_m est de mesure nulle. L'union d'une famille dénombrable de parties de mesure nulle étant elle-même de mesure nulle, S est de mesure nulle : le théorème est démontré. □

Pour clore cette section, nous faisons deux remarques. Le théorème ci-dessus est vrai pour les fonctions de plusieurs variables, avec la même preuve. Enfin, il est aussi vrai si f est définie et intégrable sur $\mathbf{R}$ en entier. En effet, $\mathbf{R}$ peut s'écrire comme l'union dénombrable des intervalles $[-n, n]$ et donc si f est positive d'intégrale nulle sur $\mathbf{R}$, elle l'est aussi sur chaque intervalle $[-n, n]$ et donc vérifie les hypothèses du théorème ; elle est donc nulle presque partout sur chaque $[-n, n]$, et comme une union dénombrable d'ensembles de mesure nulle est de mesure nulle, f est nulle presque partout.

This page intentionally left blank

Chapitre 8
Les théorèmes de convergence. Applications

8.1. Le lemme de Henstock

La théorie un peu plus avancée de l'intégrale repose presque entièrement sur un lemme facile à démontrer, mais d'une puissance remarquable. Pour des raisons de simplicité, nous ne ferons les démonstrations que dans le cas où l'intervalle d'intégration est compact.

Une *sous-subdivision pointée* de l'intervalle compact I est une collection finie J_k d'intervalles fermés, d'intérieurs disjoints, inclus dans I, et la donnée d'un point de marquage $t_k \in \mathrm{J}_k$ pour tout k : on ne suppose plus que I est l'union des J_k. Les définitions de finesse, de somme de Riemann... passent immédiatement aux sous-subdivisions. Si $\mathcal{S}_0 = \{(\mathrm{J}_k, t_k), k = 1, \ldots, n\}$ est une sous-subdivision pointée, on pose

$$\mathrm{D}(\mathcal{S}_0) = \bigcup_{k=1}^{n} \mathrm{J}_k.$$

Théorème 8.1 (Lemme de Henstock) — *Soient f une fonction intégrable sur* I *et $\varepsilon > 0$. Soit δ_ε une jauge telle que pour toute subdivision δ_ε-fine, on ait $|\mathrm{S}(f, \mathcal{S}) - \int f| \leqslant \varepsilon$. Alors, pour toute sous-subdivision δ_ε-fine $\mathcal{S}_0 = \{(\mathrm{J}_k, t_k), k = 1, \ldots, n\}$, on a :*

$$\left|\sum_{k=1}^{n}\left(f(t_k)|\mathrm{J}_k| - \int_{\mathrm{J}_k} f\right)\right| = \left|\mathrm{S}(f, \mathcal{S}_0) - \int_{\mathrm{D}(\mathcal{S}_0)} f\right| \leqslant \varepsilon.$$

En d'autres termes, si l'on a un contrôle sur les sommes de Riemann sur certaines subdivisions, on a le même contrôle sur les sommes de Riemann sur les sous-subdivisions : c'est une espèce d'uniformité spatiale de l'intégrale. Cela permet, dans

la pratique, de *concentrer l'étude sur des parties de l'intervalle* où se posent des problèmes quant aux sommes de Riemann de f.

Démonstration — Soient $\varepsilon > 0$ fixé et $\delta = \delta_\varepsilon$ comme dans l'énoncé. L'idée est de compléter la sous-subdivision $\mathcal{S}_0$ en une subdivision de I en entier, en prenant la sous-subdivision supplémentaire suffisamment fine pour que ses contributions dans les sommes de Riemann soit très petites. Le complémentaire dans I de $\mathrm{D}(\mathcal{S}_0)$ est l'union d'un nombre fini d'intervalles ouverts dont nous notons $\mathrm{K}_1, \ldots, \mathrm{K}_m$ les adhérences. Dit autrement, $\mathrm{K}_1, \ldots, \mathrm{K}_m$ sont les adhérences des *composantes connexes* de $\mathrm{I} - \mathrm{D}(\mathcal{S}_0)$. Soit $\alpha > 0$ arbitrairement petit. Comme f est intégrable sur chacun des intervalles K_i, il existe une jauge δ_i sur K_i telle que, pour toute subdivision δ_i-fine $\mathcal{Q}_i$ de K_i, on ait

$$\left|\mathrm{S}(f, \mathcal{Q}_i) - \int_{\mathrm{K}_i} f\right| \leqslant \frac{\alpha}{m}.$$

Quitte à diminuer les jauges, on peut toujours supposer que $\delta_i \leqslant \delta$. Soit maintenant $\mathcal{P}^*$ la subdivision pointée de I définie par $\mathcal{P}^* = \mathcal{S}_0 \cup \mathcal{Q}_1 \cup \cdots \cup \mathcal{Q}_m$. Elle est δ-fine, et de plus

$$\mathrm{S}(f, \mathcal{P}^*) = \mathrm{S}(f, \mathcal{S}_0) + \mathrm{S}(f, \mathcal{Q}_1) + \cdots + \mathrm{S}(f, \mathcal{Q}_m)$$

et

$$\int_{\mathrm{I}} f = \int_{\mathrm{D}(\mathcal{S}_0)} f + \int_{\mathrm{K}_1} f + \cdots + \int_{\mathrm{K}_m} f.$$

Il s'ensuit que

$$\left|\mathrm{S}(f, \mathcal{S}_0) - \int_{\mathrm{D}(\mathcal{S}_0)} f\right| \leqslant \left|\mathrm{S}(f, \mathcal{P}^*) - \int_{\mathrm{I}} f\right| + \sum_{i=1}^{m} \left|\mathrm{S}(f, \mathcal{Q}_i) - \int_{\mathrm{K}_i} f\right| \leqslant \varepsilon + \alpha,$$

et comme α a été choisi arbitrairement petit, l'inégalité du lemme de Henstock en découle. □

Corollaire 8.2 — *Sous les mêmes hypothèses, on a en plus les inégalités suivantes :*

1.
$$\sum_{k=1}^{n} \left| f(t_k)|\mathrm{J}_k| - \int_{\mathrm{J}_k} f \right| \leqslant 2\varepsilon.$$

2.
$$\left| \sum_{k=1}^{n} \left(|f(t_k)||\mathrm{J}_k| - \left| \int_{\mathrm{J}_k} f \right| \right) \right| \leqslant 2\varepsilon.$$

Démonstration — Pour obtenir la première inégalité, on casse $\mathcal{S}_0$ en deux sous-subdivisions, $\mathcal{S}_0^+$ et $\mathcal{S}_0^-$, définies par le signe de $f(t_i)|\mathrm{J}_i| - \int_{\mathrm{J}_i} f$. On peut appliquer le lemme de Henstock à $\mathcal{S}_0^+$ et $\mathcal{S}_0^-$:

$$\sum_{\mathcal{S}_0^+} \left(f(t_j) - \int_{\mathrm{I}_j} f \right) \leqslant \varepsilon \quad \text{et} \quad -\sum_{\mathcal{S}_0^-} \left(f(t_j) - \int_{\mathrm{I}_j} f \right) \leqslant \varepsilon.$$

L'addition de ces deux inégalités donne la première du corollaire. La deuxième s'obtient grâce à la première et à l'inégalité triangulaire. □

Comme nous le verrons bientôt, la seconde inégalité du corollaire est intéressant même lorsqu'on l'applique à une subdivision (complète) de l'intervalle $[a,b]$. Elle requiert cependant pour sa démonstration, même dans ce cas particulier, l'usage de sous-subdivisions et donc le lemme de Henstock.

Un exemple d'application

Le lemme de Henstock est donc très simple à démontrer, mais nous allons voir qu'il a de nombreuses et profondes applications, comme par exemple la continuité des intégrales indéfinies :

Théorème 8.3 — *Si f est une fonction intégrable sur $\mathrm{I} = [a,b]$, alors ses intégrales indéfinies sont continues.*

Démonstration — Il suffit de considérer l'intégrale indéfinie s'annulant en a :

$$\psi_f(x) = \int_a^x f.$$

Soit $c \in [a,b[$, et montrons que ψ_f est continue à droite en c. La démonstration de la continuité à gauche pour $c \in]a,b]$ sera la même, quitte à faire le changement de variable $x \mapsto -x$. Fixons $\varepsilon > 0$ et soit δ une jauge comme dans les hypothèses du lemme de Henstock. Modifions la jauge afin de forcer c à être un point de marquage : soit δ' définie par

$$\delta'(x) = \min\left\{\delta(x), \frac{1}{2}|x-c|\right\},\ x \neq c, \quad \delta'(c) = \min\left\{\delta(c), \frac{\varepsilon}{|f(c)|+1}\right\}.$$

Soient $\alpha < \delta'(c)$ et $\mathcal{S}_0$ la sous-subdivision pointée constituée de la seule paire $([c, c+\alpha], c)$. Elle est δ-fine, et donc la première inégalité du lemme de Henstock nous dit que

$$\left|\alpha f(c) - \int_c^{c+\alpha} f\right| \leqslant \varepsilon.$$

Mais $\alpha \leqslant \varepsilon(|f(c)|+1)^{-1}$, et donc

$$|\psi_f(c+\alpha) - \psi_f(c)| = \left|\int_c^{c+\alpha} f\right| \leqslant |f(c)|\alpha + \varepsilon \leqslant 2\varepsilon,$$

ce qui achève la démonstration. □

A parte historique

Il est bien beau ce théorème, mais ce que l'on aimerait vraiment savoir maintenant, c'est si ψ_f est dérivable, de dérivée f. En d'autres termes, toute fonction intégrable admet-elle une primitive ? La réponse est donnée par le théorème suivant dû à Henri Lebesgue dans le cadre de l'intégrale qui porte son nom et à Arnaud Denjoy et Oskar Perron pour l'énoncé ci-dessous.

Théorème 8.4 — *Si f est intégrable sur $[a, b]$, alors son intégrale indéfinie est continue et admet presque partout une dérivée égale à f.*

Nous démontrerons ce théorème dans l'appendice à ce chapitre comme application du lemme de recouvrement de Vitali. On peut cependant ici résumer les liens entre intégration et dérivation de la manière suivante :

- si l'on dérive une intégrale indéfinie, alors on retrouve, presque partout, la fonction dont on est parti ;
- si l'on intègre une vraie dérivée, alors on retrouve la fonction qu'on a dérivée ;
- si l'on intègre une dérivée définie seulement presque partout, alors on peut très bien ne pas pouvoir retrouver la fonction de départ. L'exemple de l'escalier du diable est à ce titre éclairant (dérivée nulle presque partout, mais fonction continue surjective de $[0, 1]$ sur lui-même).

Pour avoir un théorème complet de commutation entre dérivation et intégration, il faut se restreindre à une classe de fonctions présentant des variations locales contrôlées.

8.2. Fonctions Lebesgue-intégrables

Comme nous l'avons vu, il existe des fonctions intégrables dont la valeur absolue n'est pas intégrable. En d'autres termes, comme il existe des séries semi-convergentes, il existe des intégrales semi-convergentes (qui n'existent par dans les théories de Riemann et Lebesgue). Cependant, comme dans la théorie des séries avec les séries absolument convergentes, les fonctions à valeur absolue intégrable jouent un rôle particulier de par leur stabilité dans une large classe de processus de passage à la limite. Nous rappelons donc la définition suivante :

Définition 8.1 — *Une fonction f définie sur un intervalle quelconque* I *est dite* Lebesgue-intégrable*, ou* intégrable au sens de Lebesgue *si f* et *$|f|$ sont intégrables. On dit aussi que f est absolument intégrable.*

Si f est Lebesgue-intégrable sur I alors $f^+ = \max\{f, 0\} = (f + |f|)/2$ et $f^- = \min\{f, 0\}$ sont Lebesgue-intégrables.

Cependant l'inconvénient majeur de la définition ci-dessus est qu'elle est difficilement exploitable directement lorsqu'il s'agit de faire des manipulations algébriques

comme des sommes ou des produits, pour considérer par exemple une fonction comme $\max\{f,g\} = f + (g-f)^+$. Supposons par exemple que l'on veuille démontrer que l'espace des fonctions Lebesgue-intégrables est un espace vectoriel. Si f, g sont Lebesgue-intégrables sur I, alors $f+g$ est intégrable car f et g le sont, et $|f|+|g|$ aussi. Bien sûr $|f+g| \leqslant |f|+|g|$, mais ne pourrait-on imaginer que la fonction $|f+g|$ soit tellement atroce qu'elle ne puisse être intégrable ? On aimerait bien disposer de la proposition suivante :

Proposition 8.5 — *Soient f, g deux fonctions intégrables. Si, pour presque tout $x \in \mathrm{I}$, on a $|f(x)| \leqslant g(x)$, alors f est absolument intégrable et de plus*

$$\left| \int_{\mathrm{I}} f \right| \leqslant \int_{\mathrm{I}} |f| \leqslant \int_{\mathrm{I}} g.$$

Si l'on suppose la fonction f continue sur I, le théorème de Hake permet facilement de conclure car la régularité de la fonction f permet de l'intégrer sur les sous-intervalles compacts de I. On peut donc se poser la question suivante : existe-t-il une notion de régularité adaptée aux fonctions intégrables ? La réponse est oui et est donnée par la notion de mesurabilité d'une fonction, que nous introduirons dans le dernier chapitre de ce livre. On peut cependant s'en passer grâce au lemme de Henstock qui nous donne un contrôle local sur les fonctions intégrables et se comporte ainsi comme un substitut à une notion de régularité. Il va nous permettre de démontrer la caractérisation suivante des fonctions Lebesgue-intégrables :

Théorème 8.6 — *Soient $f : [a,b] \longrightarrow \mathbf{R}$ une fonction intégrable et $\mathrm{SP}([a,b])$ l'ensemble de toutes les subdivisions pointées de $[a,b]$. Alors $|f|$ est intégrable si et seulement si la borne supérieure*

$$\sup_{\mathcal{P}=\{(\mathrm{I}_1,t_1),\ldots,(\mathrm{I}_n,t_n)\}\,\in\,\mathrm{SP}([a,b])} \sum_{k=1}^{n} \left| \int_{\mathrm{I}_k} f(t)\,\mathrm{d}t \right|$$

est finie. Dans ce cas, elle est égale à l'intégrale de $|f|$ sur $[a,b]$.

Démonstration — Si $|f|$ est intégrable alors la positivité de l'intégrale entraîne que

$$\left| \int_{x_k}^{x_{k+1}} f(t)\,\mathrm{d}t \right| \leqslant \int_{x_k}^{x_{k+1}} |f(t)|\,\mathrm{d}t$$

pour tout sous-intervalle $[x_k, x_{k+1}]$ de $[a,b]$. Mais alors pour toute subdivision pointée $\{(\mathrm{I}_1,t_1),\ldots,(\mathrm{I}_n,t_n)\}$ de $[a,b]$ on a

$$\left| \sum_{k=1}^{n} \int_{\mathrm{I}_k} f(t)\,\mathrm{d}t \right| \leqslant \sum_{k=1}^{n} \int_{\mathrm{I}_k} |f(t)|\,\mathrm{d}t = \int_a^b |f(t)|\,\mathrm{d}t\,;$$

la borne supérieure mentionnée dans l'énoncé est donc finie et majorée par l'intégrale de $|f|$ sur $[a,b]$.

Supposons réciproquement que la borne supérieure, que nous noterons S, est finie. Fixons $\varepsilon > 0$; par définition d'une borne supérieure, il existe une subdivision pointée $\{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$ de $[a, b]$ telle que

$$\mathrm{S} - \varepsilon \leqslant \sum_{k=1}^{n} \left| \int_{\mathrm{I}_k} f(t)\,\mathrm{d}t \right| \leqslant \mathrm{S}.$$

Soit maintenant δ une jauge telle que pour toute subdivision δ-fine $\mathcal{P}$ de $[a, b]$ on ait

$$\left| \mathrm{S}(\mathcal{P}, f) - \int_a^b f(t)\,\mathrm{d}t \right| < \varepsilon\,;$$

quitte à effectuer une opération de forçage des points de marquage on peut en plus supposer que toute subdivision δ-fine définit des subdivisions de chaque intervalle $\mathrm{I}_1, \ldots, \mathrm{I}_n$. Soit alors $\mathcal{P} = \{(\mathrm{J}_1, x_1), \ldots, (\mathrm{J}_\mathrm{N}, x_\mathrm{N})\}$ une subdivision δ-fine. D'après la seconde inégalité du corollaire 8.2 appliqué à la subdivision (complète) $\mathcal{P}$ on a

$$\left| \sum_{l=1}^{\mathrm{N}} \left(|f(x_l)||\mathrm{J}_l| - \left| \int_{\mathrm{J}_l} f(t)\,\mathrm{d}t \right| \right) \right| \leqslant 2\varepsilon$$

Mais par ailleurs comme la subdivision $\mathcal{P}$ induit une subdivision pour chaque intervalle I_k, $k = 1, \ldots, n$, on a, d'après la relation de Chasles :

$$\left| \int_{\mathrm{I}_k} f(t)\,\mathrm{d}t \right| = \left| \sum_{l\,:\,\mathrm{J}_l \subset \mathrm{I}_k} \int_{\mathrm{J}_l} f(t)\,\mathrm{d}t \right| \leqslant \sum_{l\,:\,\mathrm{J}_l \subset \mathrm{I}_k} \left| \int_{\mathrm{J}_l} f(t)\,\mathrm{d}t \right| \leqslant \int_{\mathrm{I}_k} |f(t)|\,\mathrm{d}t.$$

On en déduit donc que

$$\mathrm{S} - \varepsilon \leqslant \sum_{k=1}^{n} \left| \int_{\mathrm{I}_k} f(t)\,\mathrm{d}t \right| \leqslant \sum_{l=1}^{\mathrm{N}} \left| \int_{\mathrm{J}_l} f(t)\,\mathrm{d}t \right| \leqslant \mathrm{S},$$

la deuxième inégalité dans la ligne ci-dessus découlant simplement du fait que S majore la somme à sa gauche pour *toute* subdivision, y-compris $\mathcal{P}$. En combinant cela avec l'inégalité plus haut on en déduit que

$$\left| \sum_{l=1}^{\mathrm{N}} |f(x_l)||\mathrm{J}_l| - \mathrm{S} \right| \leqslant 3\varepsilon$$

et donc que $|f|$ est intégrable, d'intégrale S. Cela achève la démonstration du théorème. □

La proposition 8.5 est alors un simple corollaire du théorème 8.6 dont nous laissons la preuve au lecteur. La classe des fonctions absolument intégrables est suffisamment importante pour mériter une notation propre :

Notation : si I est un intervalle quelconque, nous noterons $\mathcal{L}(\mathrm{I})$ l'ensemble des fonctions intégrables au sens de Lebesgue sur I.

Proposition 8.7 — *L'espace* $\mathcal{L}(\mathrm{I})$ *est un espace vectoriel. De plus,* $f \in \mathcal{L}(\mathrm{I})$ *si et seulement si* $f^+ = \max\{f, 0\}$ *et* $f_- = \min\{f, 0\}$ *sont intégrables. Enfin, si* $f, g \in \mathcal{L}(\mathrm{I})$*, alors* $\min\{f, g\}$ *et* $\max\{f, g\}$ *sont aussi intégrables au sens de Lebesgue.*

Nous laissons la preuve de cette proposition en exercice, en indiquant simplement que la proposition 8.5 vue plus haut peut être utile. On peut ici faire une remarque qui simplifie la vie : si f est Lebesgue-intégrable sur un intervalle I, alors on peut prolonger f à $\mathbf{R}$ tout entier en posant $\bar{f}(x) = 0$ si $x \notin \mathrm{I}$ et $\bar{f}(x) = f(x)$ sinon. On a alors équivalence entre $f \in \mathcal{L}(\mathrm{I})$ et $\bar{f} \in \mathcal{L}(\mathbf{R})$. On pourra ainsi dans la suite se limiter au cas de fonctions intégrables sur $\mathbf{R}$ en entier, et déduire l'énoncé correspondant pour les intervalles quelconques par une simple restriction.

8.3. Théorèmes de convergence

On ne sait calculer explicitement les intégrales que d'une toute petite classe de fonctions, comme les fonctions polynomiales par exemple. On rencontre cependant constamment, provenant des mathématiques comme de la physique, des fonctions pour lesquelles un calcul explicite de l'intégrale est impossible. On a donc souvent recours à des approximations pour attaquer ces problèmes. Une méthode générale consiste, étant donnée une fonction f, à l'écrire comme la limite d'une suite (f_n) de fonctions simples puis à démontrer que la convergence de la suite (f_n) vers f est suffisamment bonne pour que l'on puisse intervertir limite et intégrale et écrire

$$\lim_{n\to+\infty} \int f_n(x)\,\mathrm{d}x = \int \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x = \int f(x)\,\mathrm{d}x.$$

Un théorème énonçant que sous certaines hypothèses l'identité ci-dessus est vraie s'appelle un *théorème de convergence*. L'intégrale de Riemann (celle définie au moyen des jauges constantes) est intéressante du fait de son contenu intuitif et de sa capacité à intégrer les fonctions continues. Elle souffre cependant d'un manque de souplesse dans ses théorèmes de convergence. Le seul théorème vraiment général permettant d'intervertir limite et intégrale dans ce cadre est le corollaire 4.4 qui réclame une *convergence uniforme* de la suite (f_n) vers la fonction f.

Cependant la convergence uniforme est trop forte pour attraper, par approximation, toutes les fonctions intégrables, en particulier celles, un peu pathologiques, qui sortent à tous bouts de champs en physique, chimie, systèmes dynamiques et autres. Supposons par exemple que nous ne sachions faire des calculs explicites que sur des polynômes. Par passage à la limite uniforme sur un intervalle compact $[a, b]$, nous n'attraperons que des fonctions continues (en fait, toutes les fonctions continues, c'est le théorème de Stone-Weierstraß). Impossible d'approcher ainsi la fonction de Dirichlet

ou l'indicatrice d'un Cantor. Enfin, les théorèmes de convergence de l'intégrale de Riemann qui vont au delà de la convergence uniforme font presque toujours l'hypothèse de l'intégrabilité de la limite. Il est du coup impossible d'étendre la classe des fonctions intégrables par passage à la limite.

L'intégrale de Lebesgue (et plus généralement l'intégrale de Kurzweil et Henstock) permet de s'affranchir de cette limitation de l'intégrale de Riemann, car elle possède des théorèmes de convergence valables sous de très faibles hypothèses. Ce sont ces théorèmes qui rendent cette théorie si féconde et utile : c'est elle que l'on utilise dans toute l'analyse moderne (équations au dérivées partielles, distributions, théorie des opérateurs, etc.) ainsi que dans toutes les branches des mathématiques où le signe « $\int$ » fait son apparition. Enfin, c'est l'absence d'hypothèse d'intégrabilité sur la fonction limite qui rend ces énoncés si puissants.

Commençons par démontrer le théorème de convergence monotone, aussi connu sous le nom de théorème de Beppo Levi.

Théorème 8.8 — *Soient* I *un intervalle quelconque,* (f_n) *une suite de fonctions définies sur* I *et* f *une fonction définie sur* I*. On suppose que*

1. *Les fonctions* f_n *sont Lebesgue-intégrables ;*
2. *Pour presque tout* $x \in \mathrm{I}$*, la suite* $(f_n(x))$ *est croissante ;*
3. *La suite* (f_n) *converge vers* f *presque partout, c'est-à-dire pour presque tout* $x \in \mathrm{I}$*, on a* $\lim_{n\to+\infty} f_n(x) = f(x)$ *;*
4. *Et la suite des intégrales* $\left(\int_{\mathrm{I}} f_n\right)$ *est majorée.*

Alors f *est Lebesgue-intégrable et*

$$\lim_{n\to+\infty} \int_{\mathrm{I}} f_n(x)\,\mathrm{d}x = \int_{\mathrm{I}} \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x = \int_{\mathrm{I}} f(x)\,\mathrm{d}x.$$

Le même énoncé est vrai en remplaçant « croissante » par « décroissante » et « majorée » par « minorée ».

Démonstration — Nous nous placerons dans le cas où $\mathrm{I} = [a, b]$ est compact, et où la suite $f_n(x)$ converge simplement partout vers $f(x)$. La suite des intégrables

$$\int_{\mathrm{I}} f_n$$

est croissante, et comme elle est majorée, elle converge. On pose

$$\mathrm{A} := \lim_{n\to+\infty} \int_{\mathrm{I}} f_n = \sup\left\{ \int_{\mathrm{I}} f_n, n \in \mathbf{N} \right\}.$$

Fixons $\varepsilon > 0$ et $r > 0$ tels que $2^r > 4/\varepsilon$ et

$$0 \leqslant \mathrm{A} - \int_{\mathrm{I}} f_r < \varepsilon.$$

Remarquons que par monotonie, ces conditions sont en fait vérifiées pour tous les entiers plus grand que r. Pour tout $n \in \mathbf{N}$, il existe une jauge δ_n telle que, pour toute subdivision δ_n-fine $\mathcal{P}$ de I, on ait

$$\left| \mathrm{S}(f_n, \mathcal{P}) - \int_{\mathrm{I}} f_n \right| \leqslant 2^{-n}.$$

La suite (f_n) converge simplement et de manière croissante vers f, donc pour tout $x \in \mathrm{I}$, il existe un entier $n(x) \geqslant r$ tel que

$$0 \leqslant f(x) - f_{n(x)}(x) \leqslant \varepsilon,$$

et l'on pose alors $\delta(x) = \delta_{n(x)}(x)$, qui est une jauge sur I, qui va nous permettre de démontrer que f est intégrable. Une fois encore, cela illustre l'extraordinaire puissance de cette notion de jauge... Soit $\mathcal{P} = \{(\mathrm{I}_k, t_k), k = 1, \ldots, l\}$ une subdivision δ-fine de I. Comme

$$\mathrm{S}(f, \mathcal{P}) - \mathrm{A} =$$
$$\sum_{k=1}^{l} \left[\big(f(t_k)|\mathrm{I}_k| - f_{n(t_k)}(t_k)|\mathrm{I}_k| \big) + \left(f_{n(t_k)}(t_k)|\mathrm{I}_k| - \int_{\mathrm{I}_k} f_{n(t_k)} \right) + \left(\int_{\mathrm{I}_k} f_{n(t_k)} - \mathrm{A} \right) \right],$$

on en déduit, grâce aux hypothèses faites et à l'inégalité triangulaire, que

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{A}| \leqslant \sum_{k=1}^{l} \varepsilon |\mathrm{I}_k| + \sum_{k=1}^{l} \left| f_{n(t_k)}(t_k)|\mathrm{I}_k| - \int_{\mathrm{I}_k} f_{n(t_k)} \right| + \left| \sum_{i=1}^{l} \int_{\mathrm{I}_k} f_{n(t_k)} - \mathrm{A} \right|.$$

Le premier terme du membre de droite est majoré par $\varepsilon(b-a)$. Pour majorer le second, nous allons utiliser le lemme de Henstock en cassant la subdivision $\mathcal{P}$ en sous-subdivisions $\mathcal{P}_p$, définies par la condition $n(t_k) = p$, $p = r, \ldots, s = \max\{n(t_1), \ldots, n(t_l)\}$. La sous-subdivision $\mathcal{P}_p$ est clairement δ_p-fine, et la première inégalité du corollaire 8.2 au lemme de Henstock nous dit que

$$\sum_{n(t_k)=p} \left| f_{n(t_k)}|\mathrm{I}_k| - \int_{\mathrm{I}_k} f_{n(t_k)} \right| \leqslant \frac{1}{2^{p-1}},$$

ce qui, sommé sur les p variant de r à s et du fait du choix fait pour r, donne

$$\sum_{k=1}^{l} \left| f_{n(t_k)}|\mathrm{I}_k| - \int_{\mathrm{I}_k} f_{n(t_k)} \right| \leqslant \varepsilon.$$

Il nous reste à estimer le troisième terme du second membre. Pour cela, on remarque que $f_r \leqslant f_{n(t_k)} \leqslant f_s$, et donc

$$\int_{\mathrm{I}} f_r \leqslant \sum_{k=1}^{l} \int_{\mathrm{I}_k} f_{n(t_k)} \leqslant \int_{\mathrm{I}} f_s.$$

Par hypothèse sur r, cela implique

$$\mathrm{A} - \varepsilon \leqslant \sum_{k=1}^{l} \int_{\mathrm{I}_k} f_{n(t_k)} \leqslant \mathrm{A},$$

et donc le troisième terme en question est lui aussi majoré par ε. Ainsi

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{A}| \leqslant (b - a + 2)\varepsilon,$$

et f est intégrable, d'intégrale $\mathrm{A} = \lim \int_{\mathrm{I}} f_n$. Pour démontrer que $|f|$ est intégrable, et donc que f est Lebesgue-intégrable, on utilise la deuxième inégalité du corollaire 8.2, et un argument analogue. □

Voilà, le plus dur est fait ! Nous pourrons, à partir de maintenant, récolter le fruit de nos semences. Commençons par démontrer deux autres résultats de convergence. Le premier, dit lemme de Fatou, est très utile dans bien des situations.

Théorème 8.9 — *Soient f_n, g des fonctions Lebesgue-intégrables telles que $g \leqslant f_n$ pour tout n et $\liminf \int_{\mathrm{I}} f_n$ soit finie. Alors $\liminf f_n$ est Lebesgue-intégrable et*

$$(\dagger) \qquad -\infty < \int_{\mathrm{I}} g \leqslant \int_{\mathrm{I}} \liminf f_n \leqslant \liminf \int_{\mathrm{I}} f_n < +\infty.$$

Démonstration — Pour tout $k \geqslant n$, posons $g_{n,k}(x) = \min\{f_n(x), f_{n+1}(x), \dots, f_k(x)\}$, et $h_n(x) = \inf\{f_i(x), i \geqslant n\}$. L'hypothèse que $f_n \geqslant g$ implique que h_n existe. Enfin, on a, par définition,

$$h_n(x) = \lim_{k \to +\infty} g_{n,k}(x).$$

La suite $(g_{n,k})_{k \geqslant n}$ est décroissante, et de plus

$$\int_{\mathrm{I}} g \leqslant \int_{\mathrm{I}} g_{n,k} \leqslant \int_{\mathrm{I}} f_n, \text{ pour tout } k \geqslant n,$$

donc, d'après le théorème de convergence monotone, la fonction h_n est Lebesgue-intégrable et, comme

$$\int_{\mathrm{I}} h_n \leqslant \int_{\mathrm{I}} f_n,$$

on a

$$\lim_{n \to +\infty} \int h_n \leqslant \liminf_{n \to +\infty} \int_{\mathrm{I}} f_n.$$

Le second membre est fini par hypothèse, donc on peut encore appliquer le théorème de convergence monotone, et l'on obtient :

$$\int_{\mathrm{I}} \liminf_{n \to +\infty} f_n = \int_{\mathrm{I}} \lim_{n \to +\infty} h_n = \lim_{n \to +\infty} \int_{\mathrm{I}} h_n,$$

ce qui permet de conclure. □

Le second théorème de convergence est la grande victoire de la théorie de Lebesgue : le théorème de convergence dominée. Une visite approfondie de ce haut lieu de l'analyse mathématique s'impose.

Théorème de convergence dominée de Lebesgue — *Soit* (f_n) *une suite de fonctions Lebesgue-intégrables sur* I, *convergeant presque partout vers une fonction* f*. On suppose qu'il existe une fonction Lebesgue-intégrable* g *telle que*

$$|f_n(x)| \leqslant g(x), \quad \textit{pour presque tout } x \in \mathrm{I}, \ \textit{et } \forall n \in \mathbf{N}.$$

Alors f *est Lebesgue-intégrable sur* I, *et*

$$\lim_{n\to+\infty} \int_{\mathrm{I}} f_n = \int_{\mathrm{I}} \lim_{n\to+\infty} f_n = \int_{\mathrm{I}} f.$$

Démonstration — On applique le lemme de Fatou aux suites de fonctions $f_n - g$ et $g - f_n$. □

Ces théorèmes nous ouvrent la porte d'un monde merveilleux : celui de la *mesure de Lebesgue de* **R**. Nous entrouvrirons cette porte dans les derniers chapitres de ce livre. Citons deux corollaires immédiats :

Corollaire 8.10 — *Soit* (f_n) *une suite de fonction Lebesgue-intégrables sur l'intervalle compact* $[a, b]$*. Si la suite* (f_n) *est uniformément bornée, c'est-à-dire il existe une constante* $\mathrm{M} > 0$ *telle que* $|f_n(x)| \leqslant \mathrm{M}$, $\forall x \in [a, b]$, $\forall n \in \mathbf{N}$, *et si* (f_n) *converge simplement vers* f*, alors* f *est Lebesgue-intégrable et*

$$\lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Corollaire 8.11 — *Soit* (f_n) *est une suite de fonctions Lebesgue-intégrables sur* $[a, b]$ *qui converge uniformément vers* f *sur* $[a, b]$*. Alors* f *est intégrable et*

$$\lim_{n\to+\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

8.4. Intégrales dépendant d'un paramètre et dérivation sous le signe somme

Les théorèmes de passage à la limite, ou de différentiation sous le signe somme, font souvent appel, dans le cadre de la théorie de Riemann, à des hypothèses d'uniformité.

Nous allons ici obtenir les mêmes résultats sous des hypothèses plus lâches de domination. Nous ne les donnerons pas dans leur plus grande généralité, car il faudrait faire appel à la notion de fonction mesurable.

Nous considérerons une fonction $(x,t) \mapsto f(x,t)$ de deux variables, définie sur un pavé $\mathrm{I} \times \mathrm{T}$, où $\mathrm{T} = [c,d]$ est borné et où I est un intervalle quelconque. La variable t sera le paramètre.

Théorème 8.12 — *En gardant les mêmes notations, soit t_0 un point de* T. *On suppose que :*

1. *Pour tout $t \neq t_0$, la fonction $x \mapsto f(x,t)$ est Lebesgue-intégrable en la variable x sur* I *;*
2. *Pour presque tout $x \in \mathrm{I}$, on a $\lim_{t\to t_0} f(x,t) = f(x,t_0)$;*
3. *Il existe une fonction g, Lebesgue-intégrable sur* I, *telle que, pour tout $t \neq t_0$ et presque tout $x \in \mathrm{I}$ on ait*
$$|f(x,t)| \leqslant g(x).$$

Alors $x \mapsto f(x,t_0)$ est Lebesgue-intégrable sur I *et*

$$\lim_{t\to t_0} \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x = \int_{\mathrm{I}} f(x,t_0)\,\mathrm{d}x.$$

Démonstration — C'est une application immédiate du théorème de convergence dominée ; si t_n est une suite d'éléments de T tendant vers t_0, on pose $f_n(x) = f(x,t_n)$ et on applique à la suite (f_n) le théorème de Lebesgue. □

Théorème 8.13 (continuité sous le signe intégral) — *Soient $\mathrm{I} = [a,b]$ et $\mathrm{T} = [c,d]$, et $f(x,t)$ une fonction définie sur $[a,b] \times [c,d]$. On suppose que :*

1. *Pour tout $t \in \mathrm{T}$, la fonction $x \mapsto f(x,t)$ est Lebesgue-intégrable en la variable x sur* I *;*
2. *Pour tout $x \in \mathrm{I}$, la fonction $t \mapsto f(x,t)$ est continue en la variable t sur* T *;*
3. *Il existe une fonction Lebesgue-intégrable g sur* I *telle que*

$$|f(x,t)| \leqslant g(x), \quad \forall (x,t) \in \mathrm{I} \times \mathrm{T}.$$

Alors la fonction
$$t \longmapsto \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x$$
est continue sur $\mathrm{T} = [c,d]$.

Démonstration — On applique le théorème 8.12 de passage à la limite, en tout point t_0 de T. □

Voici maintenant l'énoncé que nous attendons tous, à savoir celui qui permet d'intervertir l'intégration et la dérivation partielle, ou *théorème de dérivation sous le signe somme.*

Théorème 8.14 (Dérivabilité sous le signe intégral) — *Avec les mêmes notations que plus haut, on suppose que :*

1. *Pour tout* $t_0 \in \mathrm{T}$*, la fonction* $x \mapsto f(x, t_0)$ *est Lebesgue-intégrable sur* I *;*
2. *Pour tout* $x \in \mathrm{I}$*, la fonction* $t \mapsto f(x,t)$ *est dérivable, de dérivée*

$$f_t(x) = \frac{\partial f}{\partial t}(x,t)\,;$$

3. *Il existe* g*, Lebesgue-intégrable sur* I*, telle que*

$$|f_t(x)| \leqslant g(x),\ \ \forall x \in \mathrm{I} \ \ \textit{et} \ \ \forall t \in \mathrm{T}.$$

Alors pour tout t*, les fonctions*

$$x \longmapsto f(x,t) \ \ \textit{et} \ \ x \longmapsto \frac{\partial f}{\partial t}(x,t)$$

sont Lebesgue-intégrables sur $[a,b]$*, et la fonction*

$$t \longmapsto \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x$$

est dérivable sur T*, de dérivée*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathrm{I}} f(x,t)\,\mathrm{d}x = \int_{\mathrm{I}} \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x.$$

Démonstration — Fixons $t \in \mathrm{T}$ et soit t_n une suite d'éléments de $\mathrm{T} - \{t\}$ telle que $t_n \longrightarrow t$. Alors, pour tout $x \in \mathrm{I}$, on a

$$f_t(x,t) = \lim_{n\to+\infty} \frac{f(x,t_n) - f(x,t)}{t_n - t}.$$

La formule des accroissements finis nous dit par ailleurs qu'il existe s_n (dépendant de x, t et t_n) entre t_n et t tel que

$$\frac{f(x,t_n) - f(x,t)}{t_n - t} = f_t(x, s_n).$$

Posons alors $\varphi_n(x) = f_t(x, s_n)$. La suite φ_n vérifie les hypothèses du théorème de convergence dominée, et donc $x \mapsto \lim_n \varphi_n(x) = f_t(x,t)$ est intégrable, et

$$\lim_{n\to+\infty} \int_{\mathrm{I}} \varphi_n(x) = \int_{\mathrm{I}} f_t(x,t)\,\mathrm{d}x.$$

Mais

$$\int_{\mathrm{I}} \varphi_n(x)\,\mathrm{d}x = \int_{\mathrm{I}} \frac{f(x,t_n) - f(x,t)}{t_n - t}\,\mathrm{d}x = \frac{1}{t_n - t}\left(\int_{\mathrm{I}} f(x,t_n)\,\mathrm{d}x - \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x\right),$$

et le théorème est démontré. □

Nous allons finir cette partie par l'énoncé d'un théorème utile mais rarement vu dans les cours de licence, la formule de Leibniz de dérivation d'intégrale dont l'intégrande *et* les bornes dépendent d'un paramètre.

Théorème 8.15 — *Soient* $f : \mathbf{R} \times \mathrm{T} \longrightarrow \mathbf{R}$ *et* $u, v : \mathrm{T} \longrightarrow \mathbf{R}$ *des fonctions satisfaisant les hypothèses suivantes :*

1. f *est continue sur* $\mathbf{R} \times \mathrm{T}$ *;*
2. *La dérivée partielle* $(x,t) \mapsto f_t(x,t)$ *existe, et est continue sur* $\mathbf{R} \times \mathrm{T}$ *;*
3. u *et* v *sont dérivables sur* T*.*

Alors la fonction $\mathrm{G} : \mathrm{T} \longrightarrow \mathbf{R}$*, définie par*

$$\mathrm{G}(t) = \int_{u(t)}^{v(t)} f(x,t)\,\mathrm{d}x$$

existe, et est dérivable sur T*. De plus, sa dérivée est donnée par*

$$\mathrm{G}'(t) = \int_{u(t)}^{v(t)} \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x + f(v(t),t)v'(t) - f(u(t),t)\,u'(t).$$

8.5. Appendice : le théorème de dérivation de Lebesgue

Nous avons démontré plus haut, grâce au lemme de Henstock, que si f est intégrable sur l'intervalle $[a,b]$, alors son intégrale indéfinie

$$\psi_f(x) = \int_a^x f(t)\,\mathrm{d}t$$

est continue. Si de plus f est continue, on sait que ψ_f est dérivable, de dérivée égale à f. Nous allons ici démontrer que ce fait admet une généralisation remarquable, connue sous le nom de théorème de dérivation de Lebesgue, ou « grand théorème de Lebesgue ». Il existe de nombreuses preuves de ce théorème, et nous allons en donner une s'appuyant sur le lemme de recouvrement de Vitali. Nous verrons plus loin une autre démonstration, plus simple mais faisant appel à de la théorie de la mesure, valable pour les fonctions Lebesgue-intégrables.

Théorème 8.16 — *Soit f une fonction intégrable sur l'intervalle compact $[a, b]$. Il existe un ensemble de mesure nulle $\mathrm{Z} \subset [a, b]$ tel que, pour tout $x \in [a, b] - \mathrm{Z}$, la fonction ψ_f soit dérivable de dérivée $\psi'_f(x) = f(x)$ en x.*

Démonstration — Afin de simplifier les notations, notons $\psi_f = \mathrm{F}$. Nous allons démontrer que l'ensemble E_+ des points x de $[a, b]$ en lesquels F n'admet pas de dérivée à droite, ou bien en admet qui satisfait $\mathrm{F}'_+(x) \neq f(x)$, est de mesure nulle. Le même raisonnement, effectué pour les dérivées à gauche, permettra alors de conclure en posant $\mathrm{Z} = \mathrm{E}_+ \cup \mathrm{E}_-$, puisque l'union de deux ensembles de mesure nulle l'est encore.

Si $x \in [a, b]$, la définition de la dérivabilité à droite de F en x, avec $\mathrm{F}'_+(x) = f(x)$, s'écrit ainsi : pour tout $\varepsilon > 0$, il existe $s > 0$ tel que si $x < y < x + s$ alors

$$|\mathrm{F}(y) - \mathrm{F}(x) - f(x)(y - x)| < \varepsilon|y - x|.$$

La négation de cet énoncé est alors que $x \in \mathrm{E}_+$ s'il existe $\varepsilon(x) > 0$ tel que, pour tout $s > 0$, il existe un $y_{x,s} \in \,]x, x + s[$ tel que

$$(\star) \qquad |\mathrm{F}(y_{x,s}) - \mathrm{F}(x) - f(x)(y - x)| \geqslant \varepsilon(x)|y_{x,s} - x|.$$

Par un procédé typique de la théorie de la mesure et de l'intégration, nous allons démontrer que l'on peut écrire E_+ comme une union dénombrable d'ensembles E^n_+, $n \in \mathbf{N}^*$, qui tous sont de mesure nulle. Pour cela, pour tout entier $n \geqslant 1$, on pose

$$\mathrm{E}^n_+ = \left\{ x \in \mathrm{E}_+,\ \varepsilon(x) \geqslant \frac{1}{n} \right\}.$$

On a alors

$$\mathrm{E}_+ = \bigcup_{n \geqslant 1} \mathrm{E}^n_+,$$

et donc si chacun des E^n_+ est de mesure nulle, E_+ l'est aussi. Fixons $\varepsilon > 0$ et $n \geqslant 1$; soit δ une jauge telle que, pour toute subdivision δ-fine $\mathcal{P}$ de $[a, b]$, on ait

$$\left| \mathrm{S}(f, \mathcal{P}) - \int_a^b f(t)\,\mathrm{d}t \right| \leqslant \frac{\varepsilon}{n}.$$

Considérons la famille d'intervalles suivante :

$$\mathcal{F} = \big\{ [x, y_{x,s}],\ x \in \mathrm{E}^n_+, 0 < s \leqslant \delta(x)/2 \big\}.$$

C'est un recouvrement de Vitali de E^n_+ et il existe donc une famille dénombrable $(\mathrm{I}_k = [x_k, y_k])_{k \geqslant 1}$ d'éléments de $\mathcal{F}$, deux à deux disjoints, ainsi qu'un ensemble de mesure nulle N, tels que

$$\mathrm{E}^n_+ \subset \bigcup_{k \geqslant 1} \mathrm{I}_k \cup \mathrm{N}.$$

Soit $p \geqslant 1$ un entier quelconque ; la sous-subdivision pointée $\mathcal{S}_p = \{(\mathrm{I}_1, x_1), \ldots, (\mathrm{I}_p, x_p)\}$ est δ-fine par construction, et l'on peut donc appliquer le lemme de Henstock :

$$(\dagger) \qquad \sum_{k=1}^{p} \left| f(x_k)|\mathrm{I}_k| - \int_{x_k}^{y_k} f(t)\,\mathrm{d}t \right| \leqslant \frac{2\varepsilon}{n}.$$

Par ailleurs, la relation de Chasles permet d'écrire que

$$\int_{x_k}^{y_k} f(t)\,\mathrm{d}t = \int_a^{y_k} f(t)\,\mathrm{d}t - \int_a^{x_k} f(t)\,\mathrm{d}t = \mathrm{F}(y_k) - \mathrm{F}(x_k),$$

et l'on déduit donc de $(\star)$ (en se souvenant que $y_k = y_{x_k,s}$) que

$$\begin{aligned} \sum_{k=1}^{p} \left| f(x_k)|\mathrm{I}_k| - \int_{x_k}^{y_k} f(t)\,\mathrm{d}t \right| &= \sum_{k=1}^{p} |f(x_k)(y_k - x_k) - (\mathrm{F}(y_k) - \mathrm{F}(x_k))| \\ &\geqslant \frac{1}{n} \sum_{k=1}^{p} (y_k - x_k). \end{aligned}$$

En combinant avec l'inégalité $(\dagger)$, il en découle que

$$\frac{1}{n} \sum_{k=1}^{p} (y_k - x_k) = \frac{1}{n} \sum_{k=1}^{p} |\mathrm{I}_k| \leqslant \frac{2\varepsilon}{n}.$$

Comme l'entier p a été choisi arbitrairement, le raisonnement précédent démontre en fait que

$$\sum_{k=1}^{+\infty} |\mathrm{I}_k| \leqslant 2\varepsilon\,;$$

comme E_+^n est, à un ensemble de mesure nulle près, inclus dans l'union des I_k, et que ε a lui aussi été choisi arbitrairement petit, on en déduit que E_+^n est de mesure nulle, ce qui achève la démonstration du théorème. □

Chapitre 9
Séries de Fourier

Un certain nombre d'équations fondamentales de la physique sont des équations aux dérivées partielles, dont l'inconnue est une fonction de plusieurs variables et dont les termes font intervenir des dérivées partielles de ladite inconnue. Nous ne citerons que les trois plus importantes de ces équations, les équations des ondes, de la chaleur et de Laplace.

L'équation de Laplace se rencontre lorsque l'on étudie certains systèmes au repos, ou dans un état stationnaire. C'est ainsi l'équation satisfaite par le potentiel électrique dans le vide. Si u est une fonction de d variables $(x_1, \ldots, x_d)$ (souvent $d = 2$ ou $d = 3$), on dit qu'elle satisfait à l'équation de Laplace si

$$\sum_{i=1}^{d} \frac{\partial^2 u}{\partial x_i^2} = 0.$$

On nomme la solution d'une telle équation une *fonction harmonique.* L'opérateur qui à u associe la somme des dérivées secondes de u par rapport à chacune de ses variables s'appelle le *laplacien* et se note Δ, de sorte que l'équation de Laplace s'écrit de manière condensée sous la forme

$$\Delta u = 0.$$

Il va sans dire que l'on cherche une solution de cette équation sur un domaine Ω qui est un ouvert de $\mathbf{R}^d$, et que les équations de la physique imposent alors à u de satisfaire certaines *conditions aux limites* qui s'expriment, par exemple, en imposant à u d'être définie au bord de Ω et d'y être égale à une certaine fonction donnée $f : \partial\Omega \longrightarrow \mathbf{R}$.

Une autre équation importante est l'équation de la chaleur, qui est l'équation satisfaite par la température lorsque l'on cherche à savoir comment *diffuse* la chaleur dans un milieu idéalisé pour les besoins de la cause. Son inconnue est une fonction u de $d+1$ variables, les d premières $x_1, \ldots, x_d$ représentant des variables d'espace et la dernière t une variable temporelle. Elle s'écrit

$$\frac{\partial u}{\partial t} = a^2 \Delta u.$$

Pour les mêmes raisons que pour l'équation de Laplace, l'on recherche souvent les solutions de cette équation dans un domaine de la forme $\Omega \times \mathrm{I}$ où Ω est un ouvert de $\mathbf{R}^d$ et I un intervalle de $\mathbf{R}$, en imposant des conditions aux limites.

L'équation des ondes enfin, comme son nom l'indique, est censée déterminer la propagation des ondes dans un milieu idéalisé, et elle s'écrit, de manière assez générale, sous la forme

$$\frac{\partial^2 u}{\partial t^2} = v^2 \Delta u.$$

Nous allons étudier un peu plus précisément cette dernière équation dans le cas particulier d'une *corde vibrante.* Mais avant de nous lancer, notons un caractère commun à toutes ces équations : elles sont *linéaires.* Ce que l'on entend par là c'est que si u_1 et u_2 sont deux solutions d'une de ces équations, l'équation des ondes par exemple et si μ_1, μ_2 sont deux nombres réels, alors $\mu_1 u_1 + \mu_2 u_2$ est aussi solution de la même équation. On dit aussi que ces équations satisfont au *principe de superposition.* Il vient alors une idée simple : peut-on trouver ce qui pourrait être une *base* de l'ensemble des solutions, base si possible constituée de solutions simples de ces équations ? C'est tout l'objet de la théorie des séries et de la transformée de Fourier. Le problème est qu'il faudra donner au mot « base » un sens quelque peu différent de celui couramment admis en algèbre linéaire, en admettant des sommes infinies, c'est-à-dire des sommes de *séries de fonctions.*

9.1. L'équation des cordes vibrantes

On considère une corde élastique dont les deux extrémités sont fixées et que l'on excite à un moment donné en la pinçant par exemple, et qu'on laisse ensuite évoluer librement. On suppose de plus, pour simplifier, que la corde se meut dans un plan. Rapportons ce plan à un repère et identifions-le à $\mathbf{R}^2$. La corde est alors attachée aux points $(0, 0)$ et $(\ell, 0)$, où ℓ est la longueur de la corde au repos. Le déplacement de la corde est alors modélisé par une fonction de deux variables $u(x, t)$, où $x \in [0, \ell]$ et $t \in \mathbf{R}$, de sorte qu'au temps t le point de la corde d'abscisse x est situé en $(x, u(x, t))$: $u(x, t)$ représente le *déplacement transverse* de la corde par rapport à sa position de repos.

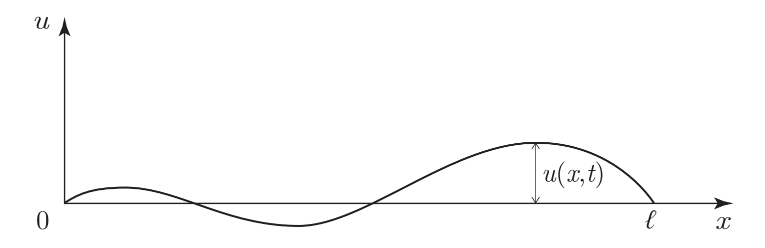


Figure 9.1 – Une corde vibrante

Nous ne nous intéresserons qu'aux petites oscillations de la corde autour de sa position d'équilibre, de sorte que l'intensité de la tension en chaque point pourra être considérée comme constante et égale à T. Nous supposerons également que la corde ne se plie pas trop violemment, ce qui se traduit par le fait que la pente (ou l'inclinaison) de la corde, donnée par la dérivée partielle

$$\frac{\partial u}{\partial x}(x,t)$$

en l'abscisse x et au temps t est petite, permettant d'identifier l'angle φ entre la tangente à la corde et l'axe des abscisses au sinus de cet angle et à la tangente à ce même angle : $\tan\varphi \approx \sin\varphi \approx \varphi$. Cela a pour autre conséquence que, si l'on considère que la corde est de masse linéique uniforme ρ, alors la masse du petit élément de corde compris entre les abscisses x et $x+\delta$ est égale à $\rho\delta$. Enfin nous négligerons les forces de frottement et de gravitation, ce qui est pertinent si la corde est fermement tendue.

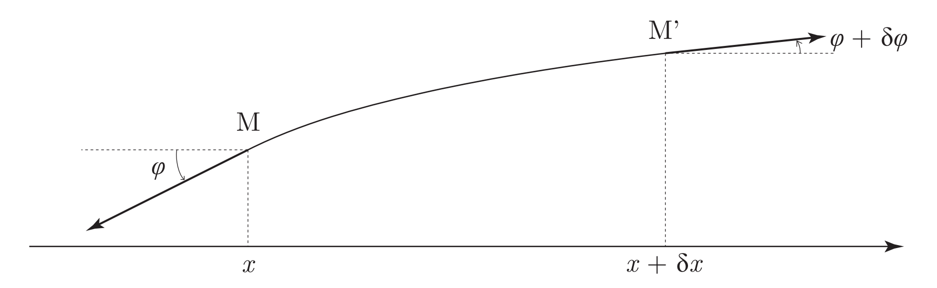


Figure 9.2 – Bilan des forces

Considérons maintenant un petit élément de corde entre les points M et M′ d'abscisses respectives x et $x + \delta x$. Soit φ l'angle de la tangente à la corde en M et $\varphi + \delta\varphi$ celui en M′. Le petit élément de corde subit deux forces de tension. La première le tirant vers la gauche et s'appliquant en M est de coordonnées $(-\mathrm{T}\cos\varphi, -\mathrm{T}\sin\varphi)$. La seconde le tirant vers la droite et s'appliquant en M′ est de coordonnées $(\mathrm{T}\cos(\varphi+\delta\varphi), \mathrm{T}\sin(\varphi+\delta\varphi))$. Dans l'approximation que nous avons faite, les composantes horizontales de ces deux forces se compensent exactement et il ne nous reste donc qu'à considérer la composante verticale de la composition de ces deux forces, qui vaut ainsi :

$$\begin{aligned}\mathrm{T}\sin(\varphi+\delta\varphi) - \mathrm{T}\sin\varphi &\approx \mathrm{T}(\tan(\varphi+\delta\varphi) - \tan\varphi)\\ &= \mathrm{T}\left(\frac{\partial u}{\partial x}(x+\delta x,t) - \frac{\partial u}{\partial x}(x,t)\right)\\ &\approx \mathrm{T}\frac{\partial^2 u}{\partial x^2}(x,t)\delta x.\end{aligned}$$

Le petit élément de corde entre M et M′ a par ailleurs une accélération verticale valant

$$\frac{\partial^2 u}{\partial t^2}(x,t)$$

et une masse égale à $\rho\delta x$, de sorte que l'équation fondamentale de la dynamique nous impose que

$$\rho\delta x\frac{\partial^2 u}{\partial t^2}(x,t) = \mathrm{T}\delta x\frac{\partial^2 u}{\partial x^2}(x,t).$$

L'évolution de l'élongation verticale u de la corde est donc gouvernée par l'équation

$$\frac{\partial^2 u}{\partial t^2} = v^2\frac{\partial^2 u}{\partial x^2},$$

où $v^2 = \mathrm{T}/\rho > 0$. La force T est homogène à des kilogrammes fois des mètres par seconde au carré. La masse linéique ρ est, elle, homogène à des kilogrammes par mètre, de sorte que v *est homogène à une vitesse.* Nous reconnaissons dans l'équation de la corde vibrante une équation des ondes en dimension $d = 1$ d'espace.

Nous allons maintenant essayer de résoudre cette équation. Il nous faut commencer par déterminer ce que nous devons connaître des conditions initiales et aux limites. Tout d'abord, la corde étant attachée en chacune de ses extrémités, nous avons

$$u(0,t) = u(\ell,t) = 0 \text{ pour tout } t.$$

L'équation faisant enfin intervenir des dérivées secondes en le temps, il nous faut connaître non seulement la position initiale de la corde mais aussi la vitesse initiale de chacun de ses points. Nous supposerons donc données deux fonctions f et g définies sur l'intervalle $[0,\ell]$ de sorte que

$$u(x,0) = f(x) \text{ et } \frac{\partial u}{\partial t}(x,0) = g(x) \text{ pour tout } x \in [0,\ell].$$

L'équation des cordes vibrantes fait intervenir séparément les dérivées par rapport au temps et à l'espace et il est naturel dans ce cas d'appliquer la méthode dite de « séparation des variables » qui consiste à chercher des solutions particulières sous la forme

$$u(x,t) = \mathrm{X}(x)\mathrm{T}(t).$$

Si une fonction mise sous cette forme est solution de l'équation des cordes vibrantes, alors

$$\frac{\mathrm{T}''}{v^2\mathrm{T}} = \frac{\mathrm{X}''}{\mathrm{X}}.$$

Dans cette équation, le premier membre ne dépend que de t tandis que le second ne dépend que de x, de sorte qu'il doit exister une constante $c \in \mathbf{R}$ telle que

$$\frac{\mathrm{T}''}{v^2\mathrm{T}} = \frac{\mathrm{X}''}{\mathrm{X}} = c.$$

Ainsi T et X satisfont aux équations différentielles ordinaires à coefficients constants

$$\mathrm{T}'' - v^2c\mathrm{T} = 0 \text{ et } \mathrm{X}'' - c\mathrm{X} = 0.$$

Si $c > 0$, les solutions sont en exponentielles réelles et ne peuvent donc s'annuler aux frontières en tout temps, de sorte que la constante c doit nécessairement être négative. Nous la noterons donc $c = -1/\lambda^2$, avec $\lambda > 0$, où λ est homogène à une longueur. Nous obtenons ainsi deux équations différentielles linéaires ordinaires

$$\mathrm{T}'' + \frac{v^2}{\lambda^2}\mathrm{T} = 0 \quad \text{et} \quad \mathrm{X}'' + \frac{1}{\lambda^2}\mathrm{X} = 0.$$

Les solutions générales de ces équations s'écrivent sous la forme

$$\mathrm{T}(t) = \mathrm{A}\cos(vt/\lambda) + \mathrm{B}\sin(vt/\lambda)$$

et

$$\mathrm{X}(x) = \mathrm{C}\cos(x/\lambda) + \mathrm{D}\sin(x/\lambda).$$

Pour simplifier quelque peu les calculs, nous allons poser $k = \lambda^{-1}$ et $\omega = v/\lambda = kv$. Exprimons le fait que $u(x,t)$ satisfait les conditions aux limites dues au fait que les extrémités de la corde sont attachées, ce qui donne une condition sur la fonction $\mathrm{X}(x)$, la fonction $\mathrm{T}(t)$ variant en permanence avec le temps :

$$\mathrm{X}(0) = 0 = \mathrm{C}, \quad \text{et} \quad \mathrm{X}(\ell) = 0 = \mathrm{C}\cos(k\ell) + \mathrm{D}\sin(k\ell).$$

Nous en déduisons que $\mathrm{C} = 0$. Si l'on cherche une solution non triviale, nous devons avoir $\mathrm{D} \neq 0$ et l'on en déduit qu'il doit exister une relation entre k et ℓ pour que $\sin(k\ell) = 0$. Ainsi doit-il exister un entier $n \in \mathbf{N}^*$ tel que $k\ell = \pi n$ ou encore

$$k = \frac{n\pi}{\ell},$$

d'où l'on tire également

$$\omega = \frac{nv\pi}{\ell}.$$

Pour chaque entier $n > 0$, nous avons trouvé une solution particulière de l'équation de la corde vibrante de la forme

$$u_n(x,t) = \sin\left(\frac{n\pi x}{\ell}\right)\left(\mathrm{A}_n\cos\left(\frac{vn\pi}{\ell}t\right) + \mathrm{B}_n\sin\left(\frac{vn\pi}{\ell}t\right)\right),$$

où A_n et B_n sont deux nombres réels arbitraires. Une telle solution élémentaire s'appelle une harmonique : elle correspond à la vibration de la corde avec une pulsation $\omega = nv\pi/\ell$, soit encore une fréquence $nv/2\ell$.

L'équation des cordes vibrantes étant *linéaire* toute combinaison linéaire des solutions u_n sera aussi une solution et, mieux, si les coefficients $\mathrm{A}_n, \mathrm{B}_n$ décroissent suffisamment vite (nous en reparlerons un peu plus loin de manière rigoureuse), nous obtenons des solutions de l'équation en considérant des *séries* de la forme

$$u(x,t) = \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{\ell}\right)\left(\mathrm{A}_n\cos\left(\frac{vn\pi}{\ell}t\right) + \mathrm{B}_n\sin\left(\frac{vn\pi}{\ell}t\right)\right).$$

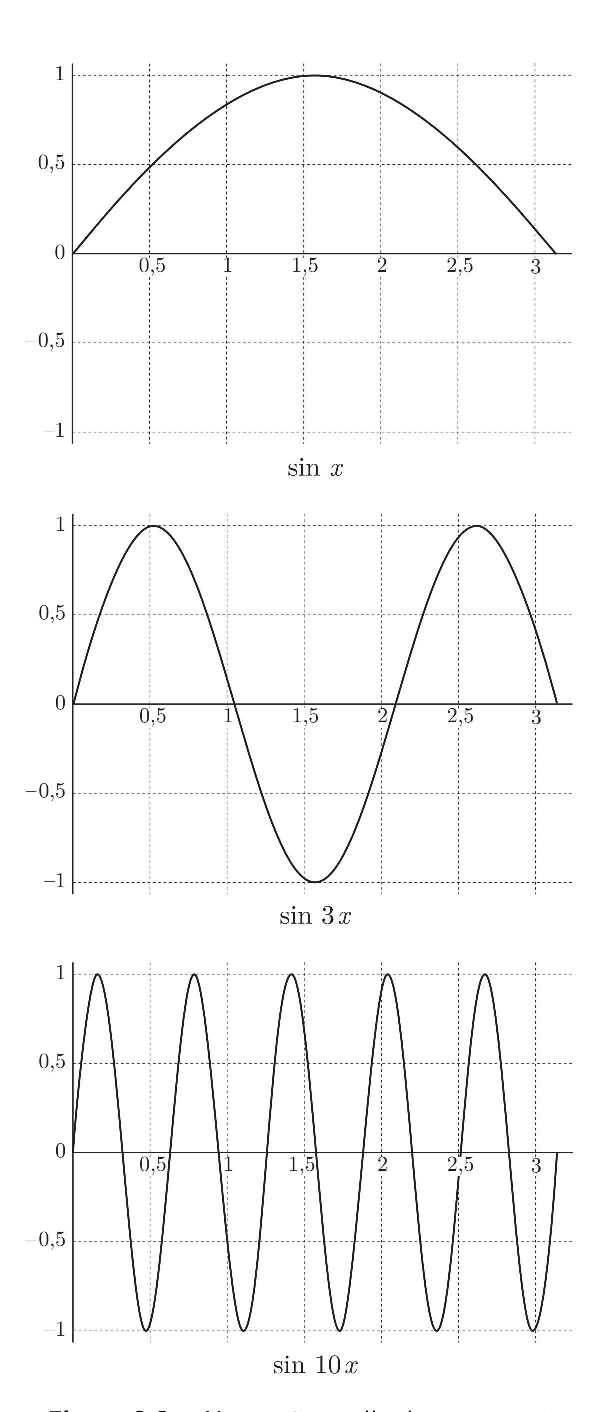


Figure 9.3 – Harmoniques d'ordre 1, 3 et 10

La question qui se pose alors naturellement est la suivante : est-ce que toute solution raisonnable de l'équation de la corde vibrante admet une écriture sous forme d'une série d'harmoniques ? Pour cela, il nous faut examiner s'il est possible, avec une telle série, de satisfaire, outre les conditions aux limites, aux conditions initiales. Rappelons que la position de la corde en $t = 0$ est donnée par la fonction f et l'on doit donc avoir, pour tout $x \in [0, \ell]$:

$$f(x) = \sum_{n=1}^{\infty} \mathrm{A}_n \sin\left(\frac{n\pi x}{\ell}\right).$$

De même, la vitesse initiale des points de la corde est donnée par la fonction g de sorte que, pour tout $x \in [0, \ell]$ nous devons avoir

$$g(x) = \sum_{n=1}^{\infty} \frac{vn\pi}{\ell} \mathrm{B}_n \sin\left(\frac{n\pi x}{\ell}\right).$$

Il nous faut donc maintenant examiner le problème suivant : une fonction f définie sur l'intervalle $[0, \ell]$, s'annulant aux extrémités de cet intervalle, peut-elle s'écrire sous la forme d'une série en sinus, comme ci-dessus ? Il va sans dire qu'il faudra sans doute se limiter à des fonctions raisonnables, par exemple continues. Ce problème est l'objet de la théorie de la décomposition des fonctions en *séries de Fourier,* que nous allons maintenant développer.

9.2. Convergence des séries de Fourier

Dans le chapitre précédent est apparu le problème de la décomposition d'une fonction sur l'intervalle $[0, \ell]$ en somme de séries de sinus de la forme $\sin(n\pi x/\ell)$. Nous allons ici donner quelques éléments de solution à ce problème qui entre dans le domaine de l'analyse connu sous le nom *d'analyse harmonique.*

Afin de simplifier les choses et ce sans perte de généralité, nous allons nous intéresser aux fonctions définies sur l'intervalle $[0, 2\pi]$. Les fonctions sinus et cosinus sont bien évidemment définies sur cet intervalle, mais aussi partout ailleurs. Cependant la valeur de ces fonctions en un réel x quelconque peut se déduire de leur valeur en un point de $[0, 2\pi]$ du fait que le sinus et le cosinus sont des *fonctions périodiques* de période 2π :

$$\cos(x + 2k\pi) = \cos(x) \text{ pour tout } x \in \mathbf{R} \text{ et tout entier } k.$$

Si f est une fonction définie sur l'intervalle $[0, 2\pi]$ il est naturel dans notre contexte d'étendre f à tout $\mathbf{R}$ en construisant une fonction 2π-périodique dont les valeurs sur $[0, 2\pi]$ sont celles de f. Nous noterons encore f la fonction ainsi étendue et nous intéresserons ici au problème de l'écriture sous forme d'une série de sinus et de cosinus d'une fonction 2π-périodique. Ce problème est double :

– étant donnée une série trigonométrique de la forme

$$\sum a_n \cos(nx) + b_n \sin(nx),$$

à quelles conditions cette série converge et, si elle converge, quelles sont les propriétés de la fonction limite ?

– étant donnée une fonction 2π-périodique f, à quelle condition peut-on l'écrire sous forme de somme d'une série trigonométrique

$$f(x) = \sum a_n \cos(nx) + b_n \sin(nx),$$

et peut-on calculer les coefficients a_n, b_n en fonction de f ?

Le premier problème est celui de la *synthèse* de signaux (fonctions) à partir de signaux sinusoïdaux simples (fonctions sinus et cosinus). Le second est celui de *l'analyse* des harmoniques simples dont est constitué un signal.

Les chapitres qui suivent vont montrer que si la fonction f est assez régulière (*grosso modo* de classe $\mathcal{C}^1$) alors f est somme d'une série trigonométrique dont les coefficients s'expriment simplement à partir de la fonction f : la *série de Fourier* de f convergera simplement et même, souvent, uniformément vers f. Si l'on assouplit le mode de convergence des séries de Fourier pour se contenter de convergence *en moyenne* (plus précisément quadratique), tout s'agence de la plus belle des manières, mais nous ne traiterons cette « théorie L^2 » des séries de Fourier qu'au chapitre 11 comme application de la construction des espaces de Lebesgue L^p.

9.2.1. Coefficients et série de Fourier d'une fonction

Supposons que l'on puisse écrire une fonction 2π-périodique f sous forme d'une série trigonométrique sous la forme (la forme particulière en $a_0/2$ du terme constant prendra sens bientôt) :

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} \big(a_n \cos(nx) + b_n \sin(nx)\big).$$

Supposons plus précisément que cette série converge suffisamment bien, par exemple uniformément, pour que l'on puisse intégrer f sur $[0, 2\pi]$ en intégrant la série au second membre terme à terme. Nous aurons alors

$$(\dagger) \qquad \int_0^{2\pi} f(x)\,\mathrm{d}x = \pi a_0 + \sum_{k=0}^{\infty} \int_0^{2\pi} (a_n \cos(nx) + b_n \sin(nx))\,\mathrm{d}x = \pi a_0.$$

Nous en déduisons donc que

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x)\,\mathrm{d}x.$$

Pour calculer les termes suivants de la série, on fait la remarque suivante : si $n \neq \pm m$ sont deux entiers relatifs quelconques, alors les fonctions $\cos(nx)\cos(mx)$, $\cos(nx)\sin(mx)$, $\cos(nx)\sin(nx)$ et $\sin(nx)\sin(mx)$ sont d'intégrale nulle entre 0 et 2π. Une intégration par parties montre par ailleurs que si $n \neq 0$ alors

$$\int_0^{2\pi} \cos^2(nx)\,\mathrm{d}x = \pi \quad \text{et} \quad \int_0^{2\pi} \sin^2(nx)\,\mathrm{d}x = \pi.$$

En calculant alors les intégrales

$$\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x \quad \text{et} \quad \int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x$$

au moyen de l'expression de f en somme de série trigonométrique, nous obtenons

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x$$

et

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x.$$

Ces coefficients sont ce que l'on appelle les *coefficients de Fourier* de la fonction f. Remarquons que pour définir ces coefficients, nous n'avons pas besoin que f soit très régulière : il suffit que f soit Lebesgue-intégrable sur $[0, 2\pi]$:

Définition 9.1 — *Soit f une fonction 2π-périodique qui soit Lebesgue-intégrable sur $[0, 2\pi]$. On appelle* coefficients de Fourier *de la fonction f les nombres a_n, b_n définis pour tout entier $n \geqslant 0$ par*

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x \quad \textit{et} \quad b_n = \int_0^{2\pi} f(x)\sin(nx)\,\mathrm{d}x.$$

La série de Fourier *de la fonction f est la série de fonctions*

$$\frac{a_0}{2} + \sum_{n\geqslant 1}\left[a_n\cos(nx) + b_n\sin(nx)\right].$$

Nous pouvons maintenant nous poser la question : l'analyse que nous avons faite ci-dessus pour trouver l'expression des coefficients de Fourier est-elle juste et mathématiquement fondée ? Plus précisément, étant donnée une fonction 2π-périodique f, sa série de Fourier converge-t-elle et, si elle converge, converge-t-elle vers f et, si elle converge vers f, converge-t-elle suffisamment bien (par exemple uniformément) pour que l'on puisse justifier l'interversion des signes somme et intégrale que nous avons effectuée dans (†) ?

La réponse à cette question est complexe. Il existe en effet des fonctions continues dont la série de Fourier diverge en un nombre dénombrable de points et des fonctions Lebesgue-intégrables dont la série de Fourier diverge *partout* (résultat dû à A. N. Kolmogoroff). Pour obtenir des théorèmes de convergence de la série de Fourier, il faut faire des hypothèses sur la *régularité* de la fonction f. La régularité d'une fonction est une mesure de la gentillesse de ses variations locales ou infinitésimales. Ainsi une fonction dérivable est plus gentille qu'une fonction simplement continue, puisqu'en plus d'avoir un graphe que l'on peut tracer sans lever le crayon de la feuille, ce graphe admet en tout point une tangente. Les problèmes de régularité sont extrêmement courants en analyse et en sont l'un des points les plus délicats. L'analyse de Fourier

est le premier domaine que l'on rencontre où ces problèmes se posent avec acuité. Les sections qui suivent vont tenter d'illustrer ce qui précède de manière concrète.

Avant de nous lancer, il est utile de transformer la série de Fourier, somme de sinus et de cosinus, en une série de Fourier dite « complexe » utilisant intelligemment la formule

$$\mathrm{e}^{ix} = \cos(x) + i\sin(x).$$

Pour tout entier $n \in \mathbf{Z}$, nous appellerons n^{e} coefficient de Fourier complexe de la fonction f le nombre

$$c_n(f) = \frac{1}{2\pi}\int_0^{2\pi} f(x)\mathrm{e}^{-inx}\,\mathrm{d}x.$$

Nous laissons le soin au lecteur tatillon de démontrer que l'on peut passer des coefficients réels a_n, b_n aux coefficients complexes c_n en toute transparence et sans douleur pour les preuves et les énoncés et qu'on a, par exemple, les formules suivantes :

$$c_0 = \frac{a_0}{2},\;\; c_n = \frac{a_n - ib_n}{2} \;\text{ si }\; n > 0 \;\text{ et }\; c_n = \frac{a_{-n} + ib_{-n}}{2} \;\text{ si }\; n < 0.$$

La série de Fourier complexe de la fonction f est alors la série

$$\sum_{n\in\mathbf{Z}} c_n(f)\mathrm{e}^{inx}.$$

Plus les coefficients c_n tendent rapidement vers 0 et meilleure sera la convergence de cette série. Si ces coefficients sont les coefficients de Fourier de la fonction f, on peut démontrer que plus la fonction f est régulière, plus ses coefficients tendent rapidement vers 0.

9.2.2. Décroissance rapide des coefficients de Fourier des fonctions régulières

Une fonction de classe $\mathcal{C}^k$ avec $k \geqslant 1$ peut, à juste titre, être qualifiée de très régulière. La proposition suivante illustre bien le lien entre régularité et décroissance des coefficients de Fourier :

Proposition 9.1 — *Si f est une fonction 2π-périodique de classe $\mathcal{C}^k$, alors la suite $(n^k c_n)_{n\in\mathbf{Z}}$ est bornée.*

Démonstration — La proposition est une conséquence simple de la formule d'intégration par parties. Si f est de classe $\mathcal{C}^k$, on peut, par k intégrations par parties, calculer le coefficient de Fourier $c_n^{(k)}$ de la fonction $f^{(k)}$, dérivée k^{e} de f :

$$c_n^{(k)} = \frac{1}{2\pi}\int_0^{2\pi} f^{(k)}(x)\mathrm{e}^{-inx}\,\mathrm{d}x = \frac{(in)^k}{2\pi}\int_0^{2\pi} f(t)\mathrm{e}^{-inx}\,\mathrm{d}x = (in)^k c_n,$$

d'où l'on déduit que

$$|n^k c_n| = |(in)^k c_n| \leqslant \frac{1}{2\pi}\int_0^{2\pi} |f^{(k)}(x)|\,\mathrm{d}x \leqslant \mathrm{M},$$

où M est un majorant quelconque de la fonction continue $|f^{(k)}|$. Cela démontre la proposition. □

Si $k \geqslant 2$, la suite (c_n) satisfait donc à $|c_n| \leqslant \mathrm{C}n^{-2}$ et l'on peut donc affirmer que la série de Fourier de f converge normalement sur l'intervalle $[0, 2\pi]$. L'analyse faite plus haut fonctionne donc parfaitement dans ce cas.

Corollaire 9.2 — *Si f est une fonction 2π-périodique de classe $\mathcal{C}^2$, sa série de Fourier converge uniformément vers f.*

Il est naturel maintenant de se poser la question : qu'en est-il des fonctions de classe $\mathcal{C}^1$? Le théorème ci-dessus énonce que les coefficients de Fourier décroissent, au mieux, comme $1/n$, ce qui n'est pas suffisant pour assurer la convergence de la série. Pour traiter ce cas il va donc nous falloir travailler un peu plus... La première étape, très générale, va consister à démontrer que les coefficients de Fourier tendent vers 0, résultat qui sera valable pour des fonctions Lebesgue-intégrables. La seconde consistera à estimer la différence entre la somme partielle de la série de Fourier et f en utilisant les propriétés remarquables du *noyau de Dirichlet.*

9.2.3. Le lemme de Lebesgue

Une fonction f est 2π-périodique si et seulement si elle est invariante par toute *translation* de sa variable par un multiple entier de 2π. De manière générale, le groupe additif $(\mathbf{R}, +)$ *agit* sur l'ensemble des fonctions définies sur $\mathbf{R}$ par translation : si f est une telle fonction et si $h \in \mathbf{R}$, on peut définir la fonction

$$\tau_h(f)(x) = f(x+h)$$

dont le graphe est obtenu à partir du graphe de f par translation de vecteur $(-h, 0)$. C'est cette action qui est au cœur de l'analyse de Fourier.

Rappelons que l'on note $\mathcal{L}(\mathbf{R})$ l'espace vectoriel des fonctions Lebesgue-intégrables sur $\mathbf{R}$. Si $f \in \mathcal{L}(\mathbf{R})$, on note

$$\|f\|_{\mathrm{L}^1} = \int_{\mathbf{R}} |f(x)|\,\mathrm{d}x$$

que l'on appelle sa (semi)-norme L^1. C'est la mesure naturelle de la taille d'une fonction Lebesgue-intégrable. La proposition suivante est facile à démontrer et nous en laissons la preuve en exercice pour le lecteur.

Proposition 9.3 — *Soit f une fonction Lebesgue-intégrable sur* $\mathbf{R}$*. L'opérateur de translation sur les fonctions satisfait aux propriétés suivantes :*

1. *Pour tout $h \in \mathbf{R}$, $\tau_h(f)$ est Lebesgue-intégrable sur $\mathbf{R}$;*
2. *Pour tout $h \in \mathbf{R}$, τ_h est une application linéaire de $\mathrm{L}^1(\mathbf{R})$ dans lui-même qui préserve la norme L^1 ;*
3. *Pour tous $h, h' \in \mathbf{R}$ on a $\tau_{h+h'}(f) = \tau_h(\tau_{h'}(f))$;*
4. *Pour tous $h, h' \in \mathbf{R}$, on a*

$$\|\tau_{h+h'}(f) - f\|_{\mathrm{L}^1} \leqslant \|\tau_h(f) - f\|_{\mathrm{L}^1} + \|\tau_{h'}(f) - f\|_{\mathrm{L}^1}.$$

Nous allons nous intéresser aux propriétés de continuité de l'opérateur de translation τ_h en fonction de h et pour cela poser $\delta_f(h) = \|\tau_h(f) - f\|_{\mathrm{L}^1}$. On a bien évidemment $\delta_f(0) = 0$. Mais on a mieux :

Proposition 9.4 — *Pour toute fonction Lebesgue-intégrable f, la fonction δ_f est continue en sa variable h.*

Démonstration — Nous allons démontrer le lemme dans le cas où f est continue et à support compact, le cas général s'en déduisant par l'application d'un théorème de densité des fonctions continues dans l'espace $\mathcal{L}(\mathbf{R})$ que nous démontrerons dans le dernier chapitre de ce livre. L'idée de la preuve dans ce cas particulier est cependant l'idée importante car elle fait ressortir les propriétés de l'action du groupe additif sur l'espace des fonctions. Soit f une fonction continue sur $\mathbf{R}$ qui soit à support compact : nous entendons par là qu'il existe un nombre $\mathrm{R} > 0$ tel que hors de l'intervalle compact $[-\mathrm{R}, \mathrm{R}]$ la fonction f soit nulle. La fonction f est en particulier bornée et il existe un nombre $\mathrm{M} > 0$ tel que, pour tout $x \in \mathbf{R}$, on ait $|f(x)| \leqslant \mathrm{M}$.

Soit maintenant $\mathrm{L} > 0$ un nombre arbitraire ; montrons que δ_f est continue sur $[-\mathrm{L}, \mathrm{L}]$. Comme f est une fonction continue, la fonction $h \mapsto |f(x+h) - f(x)|$ est continue sur $[-\mathrm{L}, \mathrm{L}]$ et ce pour tout $x \in \mathbf{R}$. De plus, pour tout $h \in [-\mathrm{L}, \mathrm{L}]$, la fonction $x \mapsto |f(x+h) - f(x)|$ est nulle hors du segment $[-(\mathrm{R}+\mathrm{L}), \mathrm{R}+\mathrm{L}]$ et y est bornée par $2\mathrm{M}$. Si l'on pose $g(x) = 2\mathrm{M}\mathbf{1}_{[-(\mathrm{R}+\mathrm{L}),\mathrm{R}+\mathrm{L}]}(x)$ alors la fonction g est Lebesgue-intégrable sur $\mathbf{R}$ et l'on a, pour tout $x \in \mathbf{R}$ et tout $h \in [-\mathrm{L}, \mathrm{L}]$:

$$|f(x+h) - f(x)| \leqslant g(x).$$

Nous pouvons donc appliquer le théorème de continuité sous le signe intégrale et en déduire que δ_f est continue sur $[-\mathrm{L}, \mathrm{L}]$. Comme L a été choisi arbitrairement, nous en déduisons que δ_f est continue sur $\mathbf{R}$. □

Nous sommes maintenant en mesure d'énoncer le lemme de Lebesgue. Pour toute fonction f Lebesgue-intégrable sur $\mathbf{R}$, on pose

$$\widehat{f}(t) = \int_{\mathbf{R}} f(x)\mathrm{e}^{itx}\,\mathrm{d}x.$$

La fonction $x \mapsto f(x)e^{itx}$ est bien Lebesgue intégrable car f l'est et qu'on la multiplie par une fonction continue et bornée sur $\mathbf{R}$. La fonction $\widehat{f}$ s'appelle la *transformée de Fourier* de la fonction f. Il est aisé de démontrer que $\widehat{f}$ est continue et bornée sur $\mathbf{R}$. Le lemme de Lebesgue énonce qu'en plus $\widehat{f}(t)$ tend vers 0 quand t tend vers l'infini :

Théorème 9.5 — *Soit f une fonction Lebesgue intégrable sur $\mathbf{R}$. Alors*

$$\lim_{t\to\pm\infty} \int_{\mathbf{R}} f(x)e^{itx}\,dx = 0,$$

autrement-dit la transformée de Fourier $t \mapsto \widehat{f}(t)$ de f tend vers zéro quand t tend vers $+\infty$ ou $-\infty$.

Démonstration — Comme $e^{i(\pi/t)t} = -1$, nous pouvons écrire, si $t \neq 0$:

$$-\widehat{f}(t) = \int_{\mathbf{R}} f(x)e^{it(x+\pi/t)}\,dx = \int_{\mathbf{R}} f\left(x - \frac{\pi}{t}\right)e^{itx}\,dx$$

et donc

$$|\widehat{f}(t)| \leqslant \frac{1}{2}\int_{\mathbf{R}} \left|f(x) - f\left(x - \frac{\pi}{t}\right)\right|\,dx.$$

La proposition 9.4 entraîne alors que $\widehat{f}(t)$ tend vers 0 quand $|t|$ tend vers l'infini, ce qui démontre le lemme de Lebesgue. □

Il pourrait sembler au lecteur que ce que nous venons de faire, traitant de fonctions Lebesgue-intégrables sur $\mathbf{R}$, n'a rien à voir avec les séries de Fourier. Il n'en est cependant rien. Soit f une fonction 2π-périodique et Lebesgue-intégrable sur $[0, 2\pi]$. La fonction $f_c(x) = \mathbf{1}_{[0,2\pi]}(x)f(x)$ est Lebesgue-intégrable sur $\mathbf{R}$ et l'on peut donc considérer sa transformée de Fourier $\widehat{f_c}$. Si n est un entier, nous avons

$$\widehat{f_c}(-n) = \int_{\mathbf{R}} f_c(x)e^{-inx}\,dx = \int_0^{2\pi} f(x)e^{-inx}\,dx$$

et nous reconnaissons le n^e coefficient de Fourier complexe de la fonction f. Nous pouvons donc énoncer le corollaire suivant du lemme de Lebesgue :

Corollaire 9.6 — *Soit f une fonction 2π-périodique et Lebesgue-intégrable sur $[0, 2\pi]$ et a_n, b_n, c_n ses coefficients de Fourier. Ils tendent vers 0 quand n tend vers l'infini :*

$$\lim_{|n|\to+\infty} c_n = 0, \quad \lim_{n\to+\infty} a_n = \lim_{n\to+\infty} b_n = 0.$$

9.2.4. Le noyau de Dirichlet

Gustav Lejeune Dirichlet est le premier mathématicien [1] à avoir démontré, en 1829, un critère général de convergence ponctuelle d'une série de Fourier. Il a fait pour cela appel à une fonction qui maintenant porte son nom : le noyau de Dirichlet.

Définition 9.2 — *Soit $n \in \mathbf{N}^*$. On appelle noyau de Dirichlet la fonction D_n définie sur $\mathbf{R}$ par*

$$\mathrm{D}_n(x) = \sum_{k=-n}^{n} \mathrm{e}^{ikx}.$$

On définit aussi le noyau de Dirichlet réel par

$$\mathrm{C}_n(x) = \frac{1}{2} + \sum_{k=1}^{n} \cos(kx).$$

En toute rigueur, le noyau de Dirichlet est la fonction des deux variables n et x. L'idée derrière l'utilisation d'un tel noyau est la suivante : si f est une fonction définie au voisinage de 0, on peut « régulariser » le comportement de f en 0 en la multipliant par D_n et en intégrant le résultat sur un intervalle centré en 0 : en effectuant ainsi une moyennisation des valeurs de f autour de 0 on obtient une fonction plus régulière. Sous de bonnes conditions, lorsque n devient grand la suite de fonctions régulières ainsi obtenues converge simplement vers f.

Nous laissons la démonstration du lemme suivant en exercice pour le lecteur :

Lemme 9.7 — *Le noyau de Dirichlet satisfait aux propriétés suivantes :*

1. Pour tout $n \in \mathbf{N}^*$ et tout $x \in \mathbf{R} - 2\pi\mathbf{Z}$,

$$\mathrm{D}_n(x) = \frac{\sin((n+1/2)x)}{\sin(x/2)}\,;$$

2. Pour tout $n \in \mathbf{N}^*$ et tout $x \in 2\pi\mathbf{Z}$,

$$\mathrm{D}_n(x) = 2n+1\,;$$

3. La fonction D_n est intégrable sur $[0, 2\pi]$ et

$$\int_0^{2\pi} \mathrm{D}_n(x)\,\mathrm{d}x = 2\pi\,;$$

1. Il faut citer ici le mathématicien Camille Deflers qui a publié en 1819 une note au *Bulletin de la Société Philomathique de Paris* où est introduit le noyau de Dirichlet. Il y démontre, du moins pour les fonctions suffisamment régulières et en utilisant une interversion entre limite et intégrale non justifiée, le lemme de Lebesgue et un théorème de convergence ponctuelle de la série de Fourier.

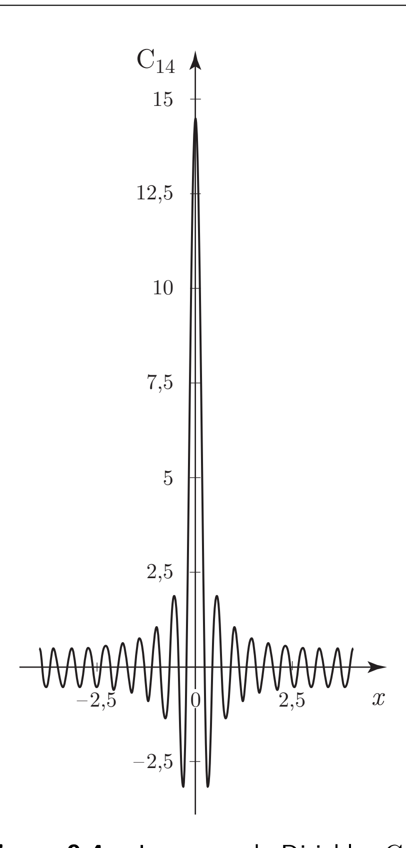


Figure 9.4 – Le noyau de Dirichlet C_{14}

4. Pour tout $n \in \mathbf{N}^*$ et tout $x \in \mathbf{R} - 2\pi\mathbf{Z}$,

$$C_n(x) = \frac{\sin((n+1/2)x)}{2\sin(x/2)}.$$

Pour exploiter le noyau de Dirichlet, il nous faut imposer à la fonction f que nous voulons étudier une certaine forme de régularité. Si f est 2π-périodique et de classe $\mathcal{C}^1$ sur $\mathbf{R}$, il est facile de démontrer que la fonction de deux variables

$$F(x,u) = \frac{1}{u}(f(x+u) - f(u))$$

est définie et *bornée* sur $\mathbf{R}^2$. Il est vite apparu qu'il était important, ne serait-ce que pour déterminer des sommes issues de l'arithmétique, d'avoir des critères de convergence ponctuelle des séries de Fourier pour des fonctions moins régulières que simplement de classe $\mathcal{C}^1$. Nous allons cependant nous en inspirer pour énoncer le théorème suivant, connu sous le nom de *critère de Dini*.

Théorème 9.8 — *Soient f une fonction 2π-périodique Lebesgue-intégrable sur $[0, 2\pi]$ et $x \in \mathbf{R}$. Si la fonction*

$$u \longmapsto \frac{f(x+u) + f(x-u) - 2f(x)}{u}$$

est Lebesgue-intégrable sur $[0, \pi]$, alors la série de Fourier de f converge au point x vers $f(x)$.

Démonstration — Notons c_n les coefficients de Fourier de f et $\mathrm{S}_n(f)$ la somme partielle de rang n de la série de Fourier de f, de sorte que

$$\mathrm{S}_n(f)(x) = \sum_{k=-n}^{n} c_k \mathrm{e}^{ikx} = \sum_{k=-n}^{n} \frac{\mathrm{e}^{ikx}}{2\pi} \int_0^{2\pi} f(t) \mathrm{e}^{-ikt} \, \mathrm{d}t$$

ce qui, en utilisant la linéarité de l'intégrale et la périodicité, nous donne

$$\mathrm{S}_n(f)(x) = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{D}_n(x-t) f(t) \, \mathrm{d}t = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{D}_n(x-t) f(t) \, \mathrm{d}t.$$

Nous en déduisons que

$$\begin{aligned} \mathrm{S}_n(f)(x) - f(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{D}_n(t)(f(x-t) - f(x)) \, \mathrm{d}t \\ &= \frac{1}{2\pi} \int_0^{\pi} \mathrm{D}_n(t)(f(x+t) + f(x-t) - 2f(x)) \, \mathrm{d}t. \end{aligned}$$

D'après le lemme 9.7, nous avons donc

$$\mathrm{S}_n(f)(x) - f(x) = \frac{1}{2\pi} \frac{f(x+t) + f(x-t) - 2f(x)}{\sin(t/2)} \sin((n+1/2)t) \, \mathrm{d}t.$$

Comme $2\sin(t/2) \approx t$ au voisinage de 0, l'hypothèse du théorème nous assure que

$$u \longmapsto \frac{f(x+u) + f(x-u) - 2f(x)}{\sin(u/2)}$$

est Lebesgue-intégrable sur $[0, \pi]$ et nous pouvons lui appliquer le lemme de Lebesgue. Ainsi $\mathrm{S}_n(f)(x) - f(x)$ tend vers 0 quand n tend vers l'infini, ce qui achève la démonstration. □

En utilisant la même stratégie de preuve, nous pouvons donner du critère de Dini une version légèrement plus précise :

Théorème 9.9 — *Soient f une fonction 2π-périodique Lebesgue-intégrable sur $[0, 2\pi]$. On suppose que f admet en tout $x \in \mathbf{R}$ des limites à gauche et à droite que nous*

noterons $f(x^-)$ et $f(x^+)$. S'il existe une fonction g Lebesgue intégrable sur $[0,\pi]$ telle que, pour tout $u \in [0,\pi]$ et presque tout x on ait

$$\left|\frac{f(x+u)+f(x-u)-(f(x^-)+f(x^+))}{u}\right| \leqslant g(x),$$

alors la série de Fourier de f converge simplement vers la fonction

$$x \longmapsto \frac{1}{2}\left[f(x^-)+f(x^+)\right].$$

Un examen minutieux des preuves faites jusqu'à maintenant permet de raffiner les théorèmes de convergence simple en des théorèmes de convergence uniforme ou normale. Nous les admettrons.

Théorème 9.10 — *Si f est 2π-périodique et de classe $\mathcal{C}^1$ par morceaux, alors sur tout intervalle de continuité de f la série de Fourier de f converge uniformément vers f.*

En guise de corollaire, nous obtenons la convergence uniforme de la série de Fourier d'une fonction de classe $\mathcal{C}^1$, répondant à la question posée plus haut. Pour les fonctions continues, une hypothèse d'absolue convergence de la série des coefficients permet également de conclure.

Théorème 9.11 — *Si f est continue et si ses coefficients de Fourier forment une série absolument convergente, c'est-à-dire si*

$$\sum_{n\in\mathbf{Z}} |c_n(f)| < +\infty,$$

alors la série de Fourier converge absolument et uniformément vers f sur $\mathbf{R}$.

Nous n'utiliserons pas ces deux derniers théorèmes dans l'application arithmétique que nous allons maintenant développer. Nous donnons une démonstration du second dans l'appendice de ce chapitre, et en donnerons une du premier dans le dernier chapitre de ce livre, comme application de la théorie L^2 des séries de Fourier.

9.2.5. Application : calcul d'une somme de nature arithmétique

Soit f la fonction 2π-périodique telle que, pour tout $x \in]-\pi,\pi]$, on ait $f(x)=|x|$. C'est une fonction continue et de classe $\mathcal{C}^1$ par morceaux qui satisfait en tout nombre x au critère de Dini. Il est aisé de calculer ses coefficients de Fourier réels a_n et b_n. Comme f est paire, ses coefficients b_n sont tous nuls. Enfin, si $n \in \mathbf{N}$ on a

$$a_n = \frac{2}{\pi}\int_0^{\pi} x\cos(nx)\,\mathrm{d}x$$

et donc $a_0 = \pi$ tandis que $a_n = 0$ si n est pair. Si $n = 2k+1$ est impair, nous avons

$$a_{2k+1} = \frac{-4}{(2k+1)^2\pi}.$$

Le critère de Dini nous permet de dire que pour tout x tel que $|x| \leqslant \pi$ on a

$$\frac{\pi}{2} - \frac{4}{\pi}\sum_{k=0}^{\infty} \frac{\cos((2k+1)x)}{(2k+1)^2} = |x|.$$

En prenant $x = 0$, nous obtenons

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}.$$

Cela va nous permettre de calculer la sommes des inverses des carrés des entiers. En effet, en scindant les entiers en pairs et impairs, nous avons

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \sum_{k=1}^{\infty} \frac{1}{4k^2} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{1}{4}\sum_{n=1}^{\infty} \frac{1}{n^2} + \sum_{k=0}^{\infty} \frac{1}{(2k+1)^2},$$

d'où l'on tire

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6},$$

ce que l'on écrit parfois sous la forme condensée

$$\zeta(2) = \frac{\pi^2}{6},$$

où ζ est la fonction *zêta* de Riemann définie par

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

De la transcendance de π évoquée il y a de cela quelques chapitres, nous déduisons que π^2 est irrationnel et donc que $\zeta(2)$ est irrationnel. De manière générale, on sait depuis Euler que $\zeta(2k)$ est irrationnel pour tout entier $k \in \mathbf{N}^*$. Le fait que $\zeta(3)$ est irrationnel est beaucoup plus difficile à démontrer. C'est un résultat dû au mathématicien Roger Apéry en 1978. On ne sait toujours pas si la fonction ζ, de la plus haute importance en arithmétique car intimement liée à la distribution des nombres premiers, prend des valeurs irrationnelles en tous les entiers impairs.

9.3. Appendice : convergence uniforme de la série de Fourier d'une fonction continue

Comme nous l'avons dit, il existe des exemples de fonctions continues dont la série de Fourier diverge en certains points. Nous allons ici démontrer le théorème suivant,

qui montre que sous une hypothèse assez facile à vérifier dans la pratique, la série de Fourier d'une fonction continue peut converger uniformément.

Théorème 9.12 — *Soient f une fonction continue 2π-périodique et c_n ses coefficients de Fourier. Si*

$$\sum_{n=-\infty}^{+\infty} |c_n| < +\infty$$

alors la série de Fourier de f converge uniformément vers f sur $\mathbf{R}$.

Démonstration — Nous pouvons, de par la périodicité de f, nous restreindre à l'intervalle $[0, 2\pi]$. L'idée de la preuve consiste à étendre les termes de la forme e^{ikx} dans la série de Fourier, qui sont des nombres complexes de module 1, en des termes de la forme $z = r\mathrm{e}^{ix}$ avec $0 \leqslant r \leqslant 1$. Si $r < 1$ le lemme de Lebesgue entraîne que la série

$$\sum_{n\in\mathbf{Z}} c_n z^n = \sum_{n\in\mathbf{Z}} c_n r^n \mathrm{e}^{inx}$$

converge normalement. L'on s'arrange ensuite pour pouvoir faire tendre r vers 1 de manière à conserver la convergence vers $f(x)$. L'objet central va, encore une fois, être un noyau : le noyau de Poisson, dont le rôle en géométrie conforme, en analyse complexe ou en théorie des probabilités est de tout premier plan.

Définition 9.3 — *On appelle noyau de Poisson la fonction définie par*

$$\mathrm{P}(r, x) = \sum_{n\in\mathbf{Z}} r^{|n|} \mathrm{e}^{inx}$$

pour $0 \leqslant r < 1$ et $x \in \mathbf{R}$.

La série donnant le noyau de Poisson est clairement normalement convergente sur $[0, 1-\varepsilon[\times \mathbf{R}$ pour tout $\varepsilon > 0$ et ainsi le noyau de Poisson est une fonction de classe $\mathcal{C}^\infty$ en ses deux variables, 2π-périodique en sa variable x. Un simple calcul montre que

$$\mathrm{P}(r, x) = \frac{1 - r^2}{1 - 2r\cos x + r^2},$$

et donc $\mathrm{P}(r, x) > 0$ pour tout r et tout x.

On a par ailleurs, pour tout $r \in [0, 1[$ et tout entier $n \geqslant 0$:

$$\left| \sum_{k=-n}^{n} r^{|k|} \mathrm{e}^{ikx} \right| \leqslant \frac{1+r}{1-r} \leqslant \frac{2}{1-r}$$

et le théorème de convergence dominée de Lebesgue nous permet d'intervertir signe somme et signe intégrale, de sorte que, pour tout $r \in [0, 1[$ on ait

$$\int_0^{2\pi} \mathrm{P}(r, x)\, \mathrm{d}x = \sum_{n\in\mathbf{Z}} \int_0^{2\pi} r^{|n|} \mathrm{e}^{inx}\, \mathrm{d}x = 2\pi.$$

Le noyau de Poisson est donc toujours d'intégrale égale à 2π, mais sa masse est très inégalement répartie et se concentre autour de 0 à mesure que r s'approche de 1, comme le démontre le lemme suivant.

Lemme 9.13 — *Pour tout δ satisfaisant à $0 < \delta < \pi$, on a*

$$\lim_{r\to 1} \int_{-\pi}^{-\delta} \mathrm{P}(r,x)\,\mathrm{d}x + \int_{\delta}^{\pi} \mathrm{P}(r,x)\,\mathrm{d}x = 0.$$

Pour démontrer le lemme, remarquons que sur $[-\pi,\pi]$, la fonction cosinus admet un unique maximum en 0. De

$$\mathrm{P}(r,x) = \frac{1-r^2}{1-2r\cos x + r^2}$$

on tire que

$$|\mathrm{P}(r,x)| \leqslant \frac{1-r^2}{1-2r\cos(\delta)+r^2}$$

pour tout r et tout x tel que $\delta \leqslant |x| \leqslant \pi$. Mais le second terme de l'inégalité précédente tend vers 0 quand r tend vers 1, ce qui permet de conclure.

Continuons la démonstration du théorème en énonçant le lemme suivant, source de tout l'intérêt que l'on peut porter au noyau de Poisson.

Lemme 9.14 — *Si f est continue, les fonctions*

$$g_r(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\mathrm{P}(r,t)\,\mathrm{d}t$$

convergent uniformément sur $\mathbf{R}$ vers $f(x)$ quand r tend vers 1.

Démontrons le lemme. La fonction f étant continue et 2π-périodique, elle est majorée par une constante que nous noterons c. Nous en déduisons que pour tout $r \in [0,1[$, $x, t \in \mathbf{R}$ et $n \in \mathbf{N}^*$ on a

$$\left| \sum_{k=-n}^{n} r^{|k|} f(t)\mathrm{e}^{ik(x-t)} \right| \leqslant \frac{2c}{1-r},$$

de sorte que l'on peut appliquer le théorème de convergence dominée de Lebesgue et intervertir les signes somme et intégrale :

$$\begin{aligned}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\mathrm{P}(r,t)\,\mathrm{d}t &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\mathrm{P}(r,x-t)\,\mathrm{d}t \\ &= \frac{1}{2\pi}\sum_{n=-\infty}^{+\infty} r^{|n|}\int_{-\pi}^{\pi} f(t)\mathrm{e}^{ik(x-t)}\,\mathrm{d}t \\ &= \sum_{n=-\infty}^{+\infty} r^{|n|} c_n \mathrm{e}^{ikx}.\end{aligned}$$

La fonction f étant continue et 2π-périodique, elle est uniformément continue. Soit $\varepsilon > 0$: il existe $\delta > 0$ tel que, pour tout $x \in \mathbf{R}$, si $|t| \leqslant \delta$ on a $|f(x) - f(x-t)| \leqslant \varepsilon$. D'après le lemme 9.13, il existe un nombre $\mathrm{R} < 1$ tel que, pour tout $r \in [\mathrm{R}, 1[$ on ait

$$\int_{-\pi}^{-\delta} \mathrm{P}(r,x)\,\mathrm{d}x + \int_{\delta}^{\pi} \mathrm{P}(r,x)\,\mathrm{d}x \leqslant \varepsilon.$$

Nous pouvons achever la démonstration du lemme 9.14 en cassant les intégrales en petits bouts (et avec des notations exotiques mais bien pratiques) :

$$\begin{aligned}\left| f(x) - \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\mathrm{P}(r,t)\,\mathrm{d}t \right| &= \left| \frac{1}{2\pi}\int_{-\pi}^{\pi} (f(x) - f(x-t))\mathrm{P}(r,t)\,\mathrm{d}t \right| \\ &\leqslant \frac{1}{2\pi}\left(\int_{-\delta}^{\delta} + \int_{-\pi}^{-\delta} + \int_{\delta}^{\pi} \right) |f(x) - f(x-t)|\,\mathrm{P}(r,t)\,\mathrm{d}t \\ &\leqslant \varepsilon(1+c),\end{aligned}$$

ce qui achève la démonstration du lemme, puisque x n'intervient pas au second membre de l'inégalité.

Nous sommes maintenant en mesure de terminer la démonstration du théorème. Nous n'avons pas encore utilisé l'hypothèse d'absolue convergence de la série des coefficients de Fourier, qui n'est autre que l'hypothèse de convergence normale de la série de Fourier de f. La série de Fourier de f est donc en particulier uniformément convergente et il ne nous reste à démontrer que le fait que sa fonction somme est égale à la fonction f. Mais, d'après le lemme 9.14, f est la limite uniforme, quand r tend vers 1, des fonctions

$$g_r(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\mathrm{P}(r,t)\,\mathrm{d}t = \sum_{k=-\infty}^{+\infty} r^{|k|} c_k \mathrm{e}^{ikx}.$$

Il nous suffit donc d'estimer la différence

$$\begin{aligned}\left| \sum_{k\in\mathbf{Z}} c_k \mathrm{e}^{ikx} - \sum_{k\in\mathbf{Z}} r^{|k|} c_k \mathrm{e}^{ikx} \right| &\leqslant \sum_{k\in\mathbf{Z}} \left|1 - r^{|k|}\right| \left|c_k\right| \\ &\leqslant \sum_{|k|\leqslant n} \left|1 - r^{|k|}\right| \left|c_k\right| + \sum_{|k|>n} \left|c_k\right|.\end{aligned}$$

Fixons $\varepsilon > 0$. La série des coefficients de Fourier étant absolument convergente, il existe un entier n tel que

$$\sum_{|k|>n} \left|c_k\right| \leqslant \varepsilon.$$

Mais par ailleurs la suite des coefficients de Fourier est bornée et donc il existe un nombre $\mathrm{R} < 1$ tel que, pour tout $r \in [\mathrm{R}, 1[$ on ait

$$\sum_{|k|\leqslant n} \left|1 - r^{|k|}\right| \left|c_k\right| \leqslant \varepsilon.$$

De là on déduit que les fonctions

$$g_r(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\mathrm{P}(r,t)\,\mathrm{d}t$$

admettent pour limite uniforme, quand r tend vers 1, la fonction somme de la série de Fourier de f. Du lemme 9.14 on déduit que f est somme de sa série de Fourier, ce qui achève la démonstration du théorème. □

Terminons cet appendice par une remarque : une fonction 2π-périodique f définit une fonction u sur le cercle unité en posant

$$u(\mathrm{e}^{it}) = f(t).$$

Si l'on note c_n les coefficients de Fourier de f, le lemme de Lebesgue nous assure que la série

$$\sum_{n=0}^{+\infty}(c_n z^n + c_{-n}\overline{z}^n)$$

converge normalement sur tout disque centré en 0 et de rayon $r < 1$. Mieux, on peut dériver terme à terme autant de fois que l'on veut et l'on conserve la convergence normale, montrant que la fonction somme de cette série, que nous noterons $\tilde{u}$, est de classe $\mathcal{C}^\infty$ (et même analytique) sur le disque unité ouvert. Un simple calcul de dérivées seconde démontre alors que $\tilde{u}$ est en fait *harmonique* sur le disque unité, c'est-à-dire qu'elle satisfait l'équation de Laplace :

$$\Delta\tilde{u} = \frac{\partial^2\tilde{u}}{\partial x^2} + \frac{\partial^2\tilde{u}}{\partial y^2} = 0.$$

La démonstration du théorème ci-dessus nous dit que si f est continue et si la série de ses coefficients de Fourier converge absolument, alors $\tilde{u}$ se prolonge par continuité au bord et y vaut précisément u : nous avons ainsi résolu un problème de Cauchy bidimensionnel en résolvant l'équation de Laplace dans le disque unité avec une condition au bord imposée. Un petit peu plus de travail nous permettrait de démontrer le théorème de représentation conforme de Riemann qui assure qu'un domaine du plan délimité par une courbe fermée, simple et suffisamment régulière, peut être mis en bijection avec le disque unité par une transformation qui, en tous les points, préserve les angles !

Partie III

Intégration des fonctions de plusieurs variables réelles et espaces de Lebesgue

This page intentionally left blank

Chapitre 10

Intégration des fonctions de plusieurs variables

En mathématiques comme dans ses applications à la physique, on est souvent confronté au calcul d'intégrales de fonctions de plusieurs variables. Si l'on étudie par exemple une distribution de charges électriques donnée par une densité $\rho(x, y, z)$ dans l'espace $\mathbf{R}^3$, la charge totale du système est donnée par l'intégrale

$$\mathrm{Q} = \int_{\mathbf{R}^3} \rho(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

Nous allons voir que les idées que nous avons exposées pour la définition de l'intégrale en une variable se généralisent sans peine aux fonctions de plusieurs, donnant les mêmes propriétés élémentaires et les mêmes théorèmes de convergence. Si tout semble se présenter sous les meilleurs auspices, plusieurs problèmes vont se poser à nous. Le premier est celui du calcul explicite d'intégrales multiples, dans des cas simples. L'idée la plus simple pour le résoudre est d'essayer de *diminuer le nombre de variables* pour se ramener à des calculs d'intégrales simples, pour lesquelles on dispose du théorème fondamental du calcul différentiel et intégral. Cela peut se faire de diverses manières et nous en verrons deux dans ce chapitre. Le *théorème de Fubini* permet de ramener le calcul d'une intégrale multiple au calcul d'une successions d'intégrales simples en intégrant en quelque sorte variable par variable. La *formule de Stokes* enfin permet de ramener le calcul de l'intégrale d'une fonction sur un domaine Ω de dimension d au calcul d'une intégrale sur le bord $\partial\Omega$ du domaine, bord qui est un objet de dimension $d-1$. Nous verrons cependant que dès que le nombre de variables atteint 2, la géométrie complique les choses et va nous imposer de ne considérer que les fonctions absolument intégrables, c'est-à-dire intégrables au sens de Lebesgue, de manière à disposer de théorèmes ayant des énoncés suffisamment simples. C'est dans ce cadre également que nous énoncerons le troisième théorème important de la théorie de l'intégration en plusieurs variables, la *formule du jacobien* ou théorème de changement de variable.

Nous ne démontrerons les théorèmes de Fubini et de changement de variable que dans le cas des fonctions continues, les outils permettant de les généraliser aux fonctions

Lebesgue-intégrables étant introduits dans le dernier chapitre de ce livre. Enfin la formule de Stokes ne sera présentée que dans sa version bi-dimensionnelle, où elle porte le nom de *formule de Green-Riemann.* Se limiter à ces cas peut sembler très restrictif mais toutes les idées importantes du sujet y surgissent déjà.

Notations : dans toute la suite, $d \geqslant 1$ est un entier fixé, la dimension de l'espace $\mathbf{R}^d$ sur lequel les fonctions que nous voudrons intégrer seront définies. Nous noterons en général un point de $\mathbf{R}^d$ en gras, sous la forme $\mathbf{x} = (x_1, \dots, x_d)$.

10.1. Pavés, subdivisions, sommes de Riemann

En dimension 1, nous avons utilisé, comme brique élémentaire dans notre construction de l'intégrale, les intervalles : leur longueur (c.-à-d. leur « volume » 1-dimensionnel) est facile à calculer. En dimension plus grande, les seuls candidats raisonnables sont les *pavés.* Dans le plan $\mathbf{R}^2$, un pavé est un rectangle de côtés parallèle aux axes de coordonnées, et dans l'espace $\mathbf{R}^3$, un pavé est un parallélépipède rectangle de côtés parallèles aux axes. De manière plus générale, on adopte la définition suivante.

Définition 10.1 — *Un pavé de* $\mathbf{R}^d$ *est une partie de* $\mathbf{R}^d$ *de la forme*

$$\mathrm{P} = \mathrm{I}_1 \times \mathrm{I}_2 \times \cdots \times \mathrm{I}_d,$$

où les I_k *sont des intervalles* compacts *de* $\mathbf{R}$. *Le volume (d-dimensionnel) de* P *est donné par le produit des longueurs des côtés*

$$|\mathrm{P}| = |\mathrm{I}_1| \cdots |\mathrm{I}_d|.$$

Ainsi, si $\mathrm{I}_k = [a_k, b_k], k = 1, \dots, d$, on peut écrire P sous la forme

$$\mathrm{P} = \{(x_1, \dots, x_d) \in \mathbf{R}^d,\ a_1 \leqslant x_1 \leqslant b_1, \dots, a_d \leqslant x_d \leqslant b_d\},$$

et le volume de P vaut alors $(b_1 - a_1) \cdots (b_d - a_d)$. En dimension $d = 1$ le volume s'appelle la longueur et en dimension $d = 2$ il s'appelle l'aire.

Munis des briques élémentaires que sont les pavés, nous pouvons maintenant mettre en route la machine des jauges et des sommes de Riemann.

Définition 10.2 — *Soit* P *un pavé de* $\mathbf{R}^d$. *On appelle subdivision pointée de* P *une famille*

$$\mathcal{P} = \{(\mathrm{P}_1, \mathbf{t}_1), \dots, (\mathrm{P}_n, \mathbf{t}_n)\}$$

telle que

1. *Pour tout* i, P_i *est un pavé de* $\mathbf{R}^d$ *;*
2. *Pour tout* i, $\mathbf{t}_i \in \mathrm{P}_i$ *et* $\mathrm{P}_i \subset \mathrm{P}$ *;*

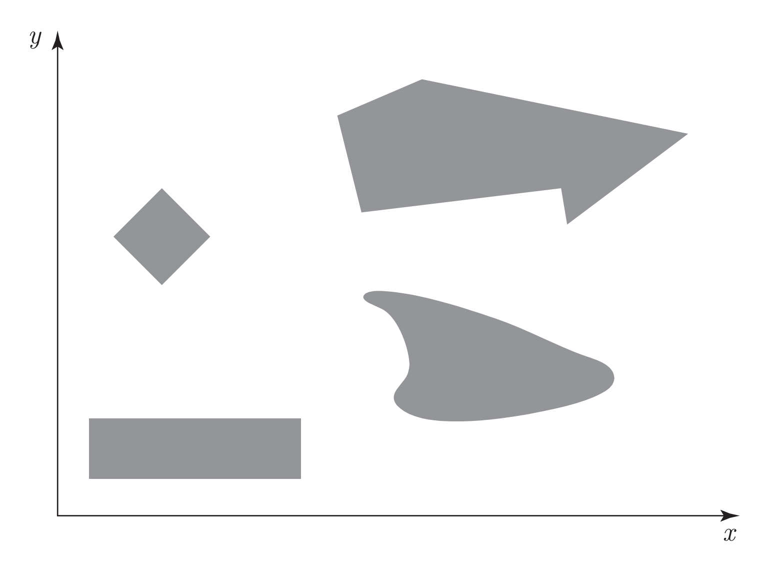


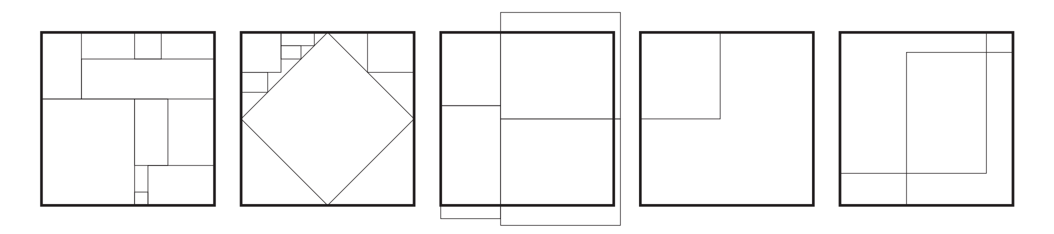
Figure 10.1 – Un pavé caché parmi trois intrus

3. Les P_i *recouvrent exactement* P *:*

$$\mathrm{P} = \bigcup_{i=1}^{n} \mathrm{P}_i \,;$$

4. Les P_i *sont d'intérieurs deux à deux disjoints : si* $i \neq j$*, alors* $\mathring{\mathrm{P}}_i \cap \mathring{\mathrm{P}}_j = \emptyset$*.*

Dans la figure qui suit, seul le premier dessin représente une subdivision du pavé unité $[0,1] \times [0,1]$ de $\mathbf{R}^2$. Les autres violent l'une ou l'autre des quatre propriétés requises.



De même que dans $\mathbf{R}$, nous allons utiliser des jauges pour contrôler la taille des pavés. Il faut commencer par remarquer que l'idée naïve de prendre le volume pour contrôler la taille est inopérante : un pavé peut être de volume petit et cependant très grand ; ainsi, le rectangle $[0, 10^{-20}] \times [0, 10^{10}]$ est de volume 10^{-10}, mais sa longueur le long de l'axe des y est de 10^{10}. Nous devons donc tenir compte de la longueur de chacun des côtés du pavé. Introduisons la notation suivante : pour $k = 1, \ldots, d$, $\pi_k : \mathbf{R}^d \longrightarrow \mathbf{R}$ est la k^{e} projection canonique, envoyant $(x_1, \ldots, x_d)$ sur x_k. Si $\mathrm{P} = \mathrm{I}_1 \times \cdots \times \mathrm{I}_d$ est un pavé, alors naturellement $\pi_k(\mathrm{P}) = \mathrm{I}_k$.

Définition 10.3 — *Une jauge sur un pavé* P *de* $\mathbf{R}^d$ *est une fonction* $\delta : \mathrm{P} \longrightarrow \mathbf{R}_+^*$. *On dit que la subdivision pointée*

$$\mathcal{P} = \{(\mathrm{P}_1, \mathbf{t}_1), \dots, (\mathrm{P}_n, \mathbf{t}_n)\}$$

est δ*-fine si, pour tout* $i = 1, \dots, n$ *et tout* $k = 1, \dots, d$, *on a*

$$\pi_k(\mathrm{P}_i) \subset \left]\pi_k(\mathbf{t}_i) - \frac{\delta(\mathbf{t}_i)}{2}, \pi_k(\mathbf{t}_i) + \frac{\delta(\mathbf{t}_i)}{2}\right[.$$

En d'autres termes, chaque pavé P_i de la subdivision est inclus dans l'hypercube centré en $\mathbf{t}_i$ de côté $\delta(\mathbf{t}_i)$. La preuve du *lemme de Cousin* pour les intervalles compacts de $\mathbf{R}$ s'étend sans modification au cas multidimensionnel, et donc toute jauge sur un pavé P de $\mathbf{R}^d$ (qui, étant fermé et borné, est compact) admet au moins une subdivision δ-fine.

Soient maintenant $f : \mathrm{P} \longrightarrow \mathbf{R}$ une fonction numérique définie sur P et $\mathcal{P} = \{(\mathrm{P}_1, \mathbf{t}_1), \dots, (\mathrm{P}_n, \mathbf{t}_n)\}$ une subdivision pointée de P ; en calquant la définition donnée pour les fonctions d'une variable, on peut parler de la somme de Riemann de f associée à $\mathcal{P}$:

$$\mathrm{S}(f, \mathcal{P}) = \sum_{i=1}^{n} f(\mathbf{t}_i)|\mathrm{P}_i|,$$

et de l'intégrabilité de la fonction f :

Définition 10.4 — *La fonction* f *est intégrable sur le pavé* P *si et seulement s'il existe un nombre* $\mathrm{S} \in \mathbf{R}$, *appelé* intégrale de f sur P, *satisfaisant à la condition suivante : pour tout* $\varepsilon > 0$, *il existe une jauge* δ *sur* P *telle que, pour toute subdivision* δ*-fine* $\mathcal{P}$ *de* P, *on ait*

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}| < \varepsilon.$$

On note alors l'intégrale de f sur P sous les diverses formes suivantes :

$$\int_{\mathrm{P}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad \text{ou} \quad \int_{\mathrm{P}} f(x_1, \dots, x_d)\,\mathrm{d}x_1 \dots \mathrm{d}x_d \quad \text{voire} \quad \int_{\mathrm{P}} f.$$

10.2. Intégration sur un ouvert quelconque

Comme nous l'avons écrit plus haut, nous allons tout de suite nous limiter à l'étude des fonctions Lebesgue-intégrables.

Définition 10.5 — *La fonction* $f : \mathrm{P} \longrightarrow \mathbf{R}$ *est dite intégrable au sens de Lebesgue, ou Lebesgue-intégrable, si* f et $|f|$ *sont intégrables sur* P. *On notera* $\mathcal{L}^1(\mathrm{P})$ *l'ensemble des fonctions Lebesgue-intégrables sur* P.

En dimension 1, il est assez naturel de se restreindre au cas des fonctions définies sur des intervalles. Cela est dû en particulier au fait que l'image d'un intervalle de $\mathbf{R}$ par un difféomorphisme est encore un intervalle (ce que l'on utilise dans la formule du changement de variable). En dimension plus grande, et dès la dimension 2, se limiter aux pavés est tout à fait insuffisant. Ainsi l'image d'un pavé par une simple application linéaire n'est plus un pavé. De même, vous pouvez essayer de démontrer qu'il existe un difféomorphisme entre le carré unité ouvert de $\mathbf{R}^2$ et le disque unité ouvert. Enfin, dans nombre de situations concrètes, les fonctions que l'on considère sont définies sur des ouverts de $\mathbf{R}^d$ de formes diverses, mais rarement des pavés. Il nous faut donc étendre la définition de fonction intégrable aux fonctions définies sur des ouverts quelconques de $\mathbf{R}^d$. Plusieurs approches de ce problème sont possibles. La plus naturelle passe par un découpage du domaine en petits bouts, puis par l'utilisation de « partitions de l'unité ». Nous prendrons ici une autre voie, plus simple techniquement, et nous nous limiterons à son énoncé, sans donner la preuve que tout cela fonctionne réellement.

Commençons par définir l'intégrabilité au sens de Lebesgue sur $\mathbf{R}^d$ tout entier.

Définition 10.6 — *La fonction $f : \mathbf{R}^d \longrightarrow \mathbf{R}$ est dite Lebesgue-intégrable (sur $\mathbf{R}^d$) si pour tout entier $n \geqslant 1$, f est Lebesgue-intégrable sur $\mathrm{P}_n = [-n, n]^d$, et si les limites*

$$\lim_{n\to+\infty} \int_{\mathrm{P}_n} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad \textit{et} \quad \lim_{n\to+\infty} \int_{\mathrm{P}_n} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x}$$

existent et sont finies. On pose alors

$$\int_{\mathbf{R}^d} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \lim_{n\to+\infty} \int_{\mathrm{P}_n} f(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Remarquons que si f n'est définie que sur le pavé P, on peut la prolonger par 0 hors de P (obtenant une fonction $\bar{f}^{\mathrm{P}}$ égale à f sur P et à 0 sur $\mathbf{R}^d - \mathrm{P}$), et f est Lebesgue-intégrable sur P si et seulement si $\bar{f}^{\mathrm{P}}$ est Lebesgue-intégrable sur $\mathbf{R}^d$.

Définition 10.7 — *Soient Ω un ouvert de $\mathbf{R}^d$ et $f : \Omega \longrightarrow \mathbf{R}$ une fonction numérique définie sur Ω. Notons $\bar{f}^{\Omega}$ la fonction définie sur $\mathbf{R}^d$ obtenue en prolongeant f par 0 hors de Ω. On dit que f est Lebesgue-intégrable sur Ω si $\bar{f}^{\Omega}$ l'est sur $\mathbf{R}^d$. On note alors*

$$\int_{\Omega} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbf{R}^d} \bar{f}^{\Omega}(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

la valeur de son intégrale et $\mathcal{L}^1(\Omega)$ l'espace des fonctions Lebesgue-intégrables sur Ω.

Munis de cette définition, nous allons maintenant pouvoir réénoncer les théorèmes vrais pour les fonctions d'une variable qui le restent en dimension supérieure.

10.3. Extension des résultats vrais dans R au cas de R^d

Beaucoup de résultats et notions que nous avons vus pour l'intégrale des fonctions d'une variable réelle restent valables pour les fonctions de plusieurs variables, et ce sans difficulté nouvelle dans leur démonstration. Le critère ce Cauchy admet le même énoncé, avec la même preuve, de même que le lemme de Henstock. Là où les choses vont coincer, c'est quand de la géométrie (formule du changement de variable) ou des notions propres au monde des plusieurs variables (formule de Fubini) vont entrer en jeu. Nous verrons ces points dans les sections suivantes.

Commençons par un fait important : les fonctions continues fournissent, sur des parties bornées, un foule d'exemple de fonctions Lebesgue-intégrables :

Théorème 10.1 — *Soient* Ω *un ouvert* borné *de* $\mathbf{R}^d$*, et* f *une fonction définie et continue sur* l'adhérence $\bar{\Omega}$ *de* Ω*. Alors* f *est Lebesgue-intégrable sur* Ω*.*

On peut généraliser cela aux fonctions continues par morceaux, même si ce concept est plus difficile à définir en plus d'une variable. Contentons-nous d'une version simplifiée :

Théorème 10.2 — *Soient* P *un pavé compact de* $\mathbf{R}^d$, $\mathcal{P} = \{\mathrm{P}_1, \ldots, \mathrm{P}_n\}$ *une subdivision de* P *en pavés et* $f : \mathrm{P} \longrightarrow \mathbf{R}$ *une fonction numérique. Alors* f *est Lebesgue-intégrable sur* P *si et seulement si* f *est Lebesgue-intégrable sur* P_i *pour tout* $i = 1, \ldots, n$*. Dans ce cas, on a de plus la formule*

$$\int_{\mathrm{P}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \sum_{i=1}^{n} \int_{\mathrm{P}_i} f(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Les inégalités de Cauchy-Schwarz, Minkowski et Hölder se généralisent sans peine au cas des fonctions continues sur l'adhérence d'un ouvert borné. La notion d'ensemble de mesure nulle admet la même définition qu'en dimension 1 :

Définition 10.8 — *Une partie* $\mathrm{N} \subset \mathbf{R}^d$ *est dite de mesure nulle si pour tout* $\varepsilon > 0$ *il existe une famille au plus dénombrable* $(\mathrm{P}_n)_{n\in\mathbf{N}}$ *de pavés de* $\mathbf{R}^d$ *tels que*

$$\mathrm{N} \subset \bigcup_{n\in\mathbf{N}} \mathrm{P}_n \quad \textit{et} \quad \sum_{i\in\mathbf{N}} |\mathrm{P}_n| < \varepsilon.$$

Les propriétés élémentaires des parties de mesure nulle sont les mêmes que dans $\mathbf{R}$: l'ensemble vide, les points, les ensembles dénombrables, les unions dénombrables de parties de mesures nulles sont de mesure nulle et toute partie d'une partie de mesure nulle est de mesure nulle. On dit qu'une propriété est vraie presque partout si elle est vraie hors d'un ensemble de mesure nulle. Mais il est une classe nouvelle importante d'ensembles de mesure nulle : *une hypersurface de classe* $\mathcal{C}^1$ *de* $\mathbf{R}^d$ *est de mesure nulle.* Ainsi une courbe dans $\mathbf{R}^2$, une surface dans $\mathbf{R}^3$ sont des objets de mesure nulle.

Pour voir pourquoi cela est vrai, considérons le disque $\mathrm{D} = \{x^2 + y^2 \leqslant 1, z = 0\}$ dans l'espace $\mathbf{R}^3$ muni des coordonnées (x, y, z). On peut dans ce cas recouvrir D par un pavé P dont la base sur le plan (x, y) contient D, et ensuite ajuster son volume en prenant une extension dans la direction verticale (z) aussi petite que l'on veut. On peut prendre par exemple $\mathrm{P} = [-1,1] \times [-1,1] \times [-\varepsilon/8, \varepsilon/8]$; on a alors $\mathrm{D} \subset \mathrm{P}$ et $|\mathrm{P}| \leqslant \varepsilon$. Il est heureux qu'il en soit ainsi, puisque cela permet de glisser un disque compact dans une pochette de faible volume : pas besoin d'avoir un entrepôt géant pour stocker l'intégrale des œuvres enregistrées de Frank Zappa ou Jean-Sébastien Bach.

Les ensembles de mesure nulle sont les ensembles négligeables dans la théorie de l'intégration, comme le résume le théorème suivant :

Théorème 10.3 — *Soient Ω un ouvert de $\mathbf{R}^d$ et f une fonction définie sur Ω.*

1. *Si f est nulle presque partout, alors f est Lebesgue-intégrable sur Ω et son intégrale est nulle.*
2. *Si f est positive et Lebesgue-intégrable, alors*

$$\int_\Omega f(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0$$

si et seulement si $f(\mathbf{x}) = 0$ presque partout.

Ainsi, on peut modifier les valeurs d'une fonction sur un ensemble de mesure nulle sans changer son caractère intégrable ni la valeur de son intégrale.

Le lemme de Henstock étant vrai dans $\mathbf{R}^d$ (avec la même preuve qu'en dimension 1, légèrement plus technique), on y dispose des mêmes théorèmes de convergence que pour les fonctions d'une variable, ainsi que des théorème de continuité et dérivabilité sous le signe intégral.

Théorème 10.4 — *Soient Ω un ouvert de $\mathbf{R}^d$, (f_n) une suite de fonctions définies sur Ω et f une fonction définie sur Ω. On suppose que*

1. *Les fonctions f_n sont Lebesgue-intégrables ;*
2. *Pour presque tout $\mathbf{x} \in \Omega$, la suite $(f_n(\mathbf{x}))$ est croissante ;*
3. *La suite (f_n) converge presque partout vers f, c'est-à-dire pour presque tout $\mathbf{x} \in \Omega$ on a $\lim_{n\to+\infty} f_n(\mathbf{x}) = f(\mathbf{x})$;*
4. *La suite de nombres réels $\int_\Omega f_n(\mathbf{x})\,\mathrm{d}\mathbf{x}$ est majorée.*

Alors f est Lebesgue-intégrable et

$$\lim_{n\to+\infty} \int_\Omega f_n(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_\Omega \lim_{n\to+\infty} f_n(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_\Omega f(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Le même énoncé est vrai en remplaçant « croissante » par « décroissante » et « majorée » par « minorée ».

Le lemme de Fatou est vrai avec le même énoncé, aussi nous ne le répéterons pas ; par contre et vu son importance, réénonçons le théorème de convergence dominée de Lebesgue :

Théorème de convergence dominée de Lebesgue — *Soit (f_n) une suite de fonctions Lebesgue-intégrables sur Ω, convergeant presque partout vers une fonction f. On suppose qu'il existe une fonction g Lebesgue-intégrable sur Ω telle que*

$$|f_n(\mathbf{x})| \leqslant g(\mathbf{x}), \quad \text{pour presque tout } \mathbf{x} \in \Omega, \text{ et } \forall n \in \mathbf{N}.$$

Alors f est Lebesgue-intégrable sur Ω, et

$$\lim_{n\to+\infty} \int_{\mathrm{I}} f_n = \int_{\mathrm{I}} \lim_{n\to+\infty} f_n = \int_{\mathrm{I}} f.$$

Muni de ces théorèmes, on peut, exactement comme dans le cas des fonctions d'une variable, démontrer des théorèmes de continuité et de dérivabilité sous le signe intégral. Nous laissons le soin au lecteur consciencieux de rédiger ceux-ci dans le cadre des fonctions de plusieurs variables.

10.4. Les théorèmes de Fubini et de Tonelli

Pour l'instant, nous ne disposons pas de résultats permettant le calcul effectif d'intégrales (hormis celle des fonctions constantes sur des pavés), même pour des fonctions très simples comme des polynômes. Pour faire de tels calculs, deux approches sont possibles : les réduire à un calcul d'intégrales de fonctions d'une variable, ou essayer de trouver un analogue multidimensionnel de la formule fondamentale du calcul différentiel et intégral ; le théorème de Fubini permet la mise en place de la première possibilité, en montrant que le calcul d'une intégrale multiple peut se ramener à une suite de calculs d'intégrales simples. Nous verrons plus loin que la deuxième possibilité existe, mais requiert l'introduction de concepts nouveaux.

Avant d'énoncer le théorème de Fubini, mettons en place les protagonistes. Tout d'abord, on se donne un pavé P de $\mathbf{R}^d$, que l'on écrit sous la forme $\mathrm{P} = \mathrm{I}_1 \times \cdots \times \mathrm{I}_d$, où I_k est un intervalle compact de $\mathbf{R}$. Ensuite, on sépare les d variables en deux groupes de d' (resp. $d - d'$) variables, ce que l'on matérialise en posant $\mathrm{P}_1 = \mathrm{I}_1 \times \cdots \times \mathrm{I}_{d'}$ et $\mathrm{P}_2 = \mathrm{I}_{d'+1} \times \cdots \times \mathrm{I}_d$. Ainsi P_1 (resp. P_2) est un pavé de $\mathbf{R}^{d'}$ (resp. $\mathbf{R}^{d-d'}$). Voir la figure 10.2 illustrant cette situation pour un pavé P de $\mathbf{R}^3$, avec $d' = 2$.

Si maintenant f est une fonction réelle définie sur P, on peut l'écrire sous la forme $f(\mathbf{x}) = f(\mathbf{x}_1, \mathbf{x}_2)$, avec $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) \in \mathrm{P} = \mathrm{P}_1 \times \mathrm{P}_2$. Si $\mathbf{x}_1$ est un élément de P_1, on peut définir la fonction partielle $f(\mathbf{x}_1, \cdot) : \mathrm{P}_2 \longrightarrow \mathbf{R}$, en posant $f(\mathbf{x}_1, \cdot)(\mathbf{x}_2) = f(\mathbf{x}_1, \mathbf{x}_2)$, qui est obtenue en fixant la première variable et en faisant varier la seconde.

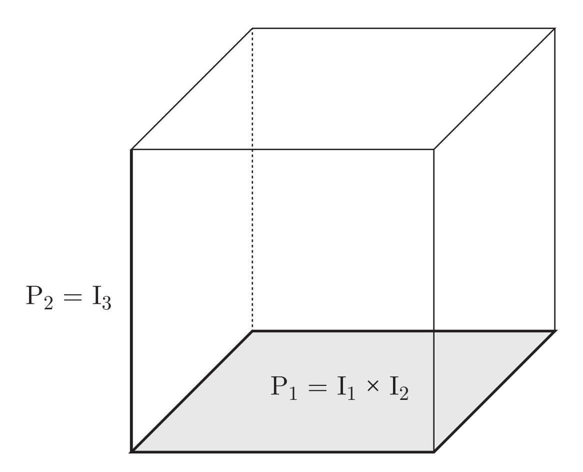


Figure 10.2 – Décomposition d'un pavé de $\mathbf{R}^3$ pour le théorème de Fubini

On a de même, pour tout $\mathbf{x}_2 \in P_2$ fixé, une fonction partielle de la première variable $f(\cdot, \mathbf{x}_2) : P_1 \longrightarrow \mathbf{R}$. Nous pouvons maintenant énoncer le théorème de Fubini :

Théorème de Fubini — *Avec les notations qui précèdent, supposons que* $f : P = P_1 \times P_2 \longrightarrow \mathbf{R}$ *soit Lebesgue-intégrable. Alors*

1. *Pour presque tout* $\mathbf{x}_1 \in P_1$*, la fonction* $f(\mathbf{x}_1, \cdot) : P_2 \longrightarrow \mathbf{R}$ *est Lebesgue-intégrable sur* P_2 *;*

2. *Pour presque tout* $\mathbf{x}_2 \in P_2$*, la fonction* $f(\cdot, \mathbf{x}_2) : P_1 \longrightarrow \mathbf{R}$ *est Lebesgue-intégrable sur* P_1 *;*

3. *Pour presque tout* $\mathbf{x}_1 \in P_1$*, la fonction*
$$\mathbf{x}_1 \longmapsto \int_{P_2} f(\mathbf{x}_1, \mathbf{x}_2)\, \mathrm{d}\mathbf{x}_2$$
est Lebesgue-intégrable sur P_2 *;*

4. *Pour presque tout* $\mathbf{x}_2 \in P_2$*, la fonction*
$$\mathbf{x}_2 \longmapsto \int_{P_1} f(\mathbf{x}_1, \mathbf{x}_2)\, \mathrm{d}\mathbf{x}_2$$
est Lebesgue-intégrable sur P_1*.*

On a de plus la formule suivante, dite « formule de Fubini » :
$$\int_{P_1 \times P_2} f(\mathbf{x})\, \mathrm{d}\mathbf{x} = \int_{P_1} \left(\int_{P_2} f(\mathbf{x}_1, \mathbf{x}_2)\, \mathrm{d}\mathbf{x}_2 \right) \mathrm{d}\mathbf{x}_1 = \int_{P_2} \left(\int_{P_1} f(\mathbf{x}_1, \mathbf{x}_2)\, \mathrm{d}\mathbf{x}_1 \right) \mathrm{d}\mathbf{x}_2.$$

Pour ramener le calcul de l'intégral de f à une suite de calculs d'intégrales unidimensionnelles, on applique itérativement le théorème de Fubini, en écrivant successivement $\mathrm{P} = \mathrm{I}_1 \times (\mathrm{I}_2 \times \cdots \times \mathrm{I}_d)$, $\mathrm{I}_2 \times \cdots \times \mathrm{I}_d = \mathrm{I}_2 \times (\mathrm{I}_3 \times \cdots \times \mathrm{I}_d)$, etc. On tombe alors sur la formule suivante :

$$\int_{\mathrm{P}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathrm{I}_1} \left(\int_{\mathrm{I}_2} \left(\cdots \left(\int_{\mathrm{I}_d} f(x_1, \ldots, x_d)\,\mathrm{d}x_d \right) \cdots \right) \mathrm{d}x_2 \right) \mathrm{d}x_1.$$

Le théorème de Tonelli, antérieur à celui de Fubini, ne s'applique qu'aux fonctions positives, mais il permet de simplifier la vérification de l'intégrabilité de f sur P.

Théorème de Tonelli — *Soit* $f : \mathrm{P} = \mathrm{P}_1 \times \mathrm{P}_2 \longrightarrow \mathbf{R}$ *une fonction* positive. *Si l'une des trois intégrales*

$$\int_{\mathrm{P}} f(\mathbf{x})\,\mathrm{d}\mathbf{x},\ \int_{\mathrm{P}_1} \left(\int_{\mathrm{P}_2} f(\mathbf{x}_1, \mathbf{x}_2)\,\mathrm{d}\mathbf{x}_2 \right) \mathrm{d}\mathbf{x}_1\ \int_{\mathrm{P}_2} \left(\int_{\mathrm{P}_1} f(\mathbf{x}_1, \mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1 \right) \mathrm{d}\mathbf{x}_2$$

existe (et est finie), alors les deux autres aussi, et les trois sont égales. En particulier, f *est Lebesgue-intégrable sur* P*.*

Nous ne donnons pas de preuve générale de ces deux théorèmes : elles ne sont pas difficiles, mais réclament de fastidieuses mises en place concernant les subdivisions pointées. Contentons-nous de prouver le théorème de Fubini dans le cas où f est à deux variables, définie sur le pavé unité $\mathrm{P} = [0,1] \times [0,1]$ et y est continue; les ingrédients géométriques de la preuve générale se trouvent tous dans ce cas simple.

Preuve du théorème de Fubini dans un cas simple — Soit donc f une fonction continue sur le pavé $\mathrm{P} = [0,1] \times [0,1]$. Comme P est compact, la fonction f est uniformément continue sur P ; nous fixons $\varepsilon > 0$ et prenons pour jauge δ la jauge constante donnée par la continuité uniforme de f. Si $\mathcal{P}$ est une subdivision δ-fine, on peut toujours la raffiner, sans changer ses sommes de Riemann, pour qu'elle s'organise comme un mur de briques (on parle de subdivision maçonnée) :

$$\mathcal{P} = \{(\mathrm{P}_{1,j}, t_{1,j}), j = 1, \ldots, n_1; \ldots; (\mathrm{P}_{k,j}, t_{k,j}), j = 1, \ldots, n_k\},$$

et l'on suppose que pour tous i, j, la projection $\pi_2(\mathrm{P}_{i,j}) = \mathrm{Q}_i$ sur l'axe vertical est un intervalle indépendant de j tel que $\{\mathrm{Q}_1, \ldots, \mathrm{Q}_k\}$ forme une subdivision de $[0,1]$. On choisit dans chaque point de Q_i un point de marquage t_i, de sorte que

$$\mathcal{P}' = \{(\mathrm{Q}_1, t_1), \ldots, (\mathrm{Q}_k, t_k)\}$$

soit une subdivision pointée de $[0,1]$ (*voir* la figure 10.3).

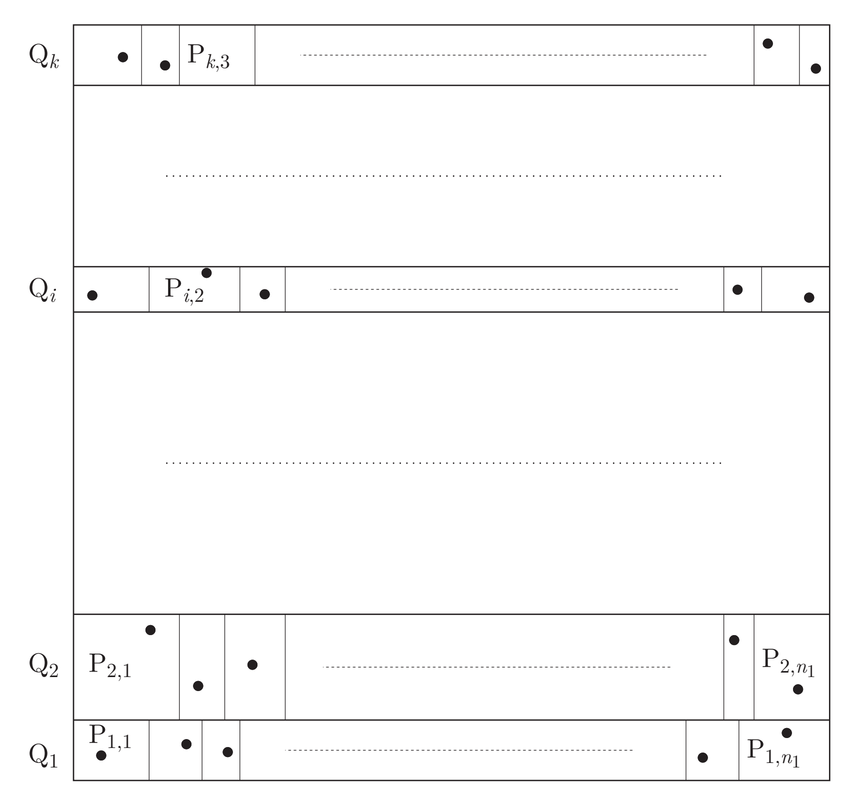


Figure 10.3 – La subdivision maçonnée $\mathcal{P}$

On utilise ensuite la continuité uniforme de f pour estimer les sommes de Riemann sur les tranches horizontales de la subdivision $\mathcal{P}$: soit $i \in \{1, \ldots, k\}$, alors

$$\begin{aligned}
&\left| \int_0^1 f(x, t_i)\, \mathrm{d}x - \sum_{j=1}^{n_i} |\mathrm{P}_{i,j}||\mathrm{Q}_i|^{-1} f(t_{i,j}) \right| \\
&\qquad = \left| \sum_{j=1}^{n_i} \left(\int_{\pi_1(\mathrm{P}_{i,j})} f(x, t_i)\, \mathrm{d}x - |\pi_1(\mathrm{P}_{i,j})| f(t_{i,j}) \right) \right| \\
&\qquad \leqslant \sum_{j=1}^{n_i} \left(\int_{\pi_1(\mathrm{P}_{i,j})} |f(x, t_i) - f(t_{i,j})|\, \mathrm{d}x \right) \\
&\qquad \leqslant \sum_{j=1}^{n_i} \varepsilon |\pi_1(\mathrm{P}_{i,j})| = \varepsilon,
\end{aligned}$$

où la dernière inégalité découle de l'uniforme continuité de f ; en effet, si $x \in \pi_1(\mathrm{P}_{i,j})$, alors $\|(x, t_i) - t_{i,j}\| < \delta$ et donc $|f(x, t_i) - f(t_{i,j})| < \varepsilon$.

Par ailleurs, comme l'application

$$t \longmapsto \int_0^1 f(x,t)\,\mathrm{d}x$$

est continue sur $[0,1]$, elle y est uniformément continue, et l'on en tire l'inégalité suivante :

$$\left| \int_0^1 \left(\int_0^1 f(x,t)\,\mathrm{d}x \right) \mathrm{d}t - \sum_{i=1}^{k} |\mathrm{Q}_i| \int_0^1 f(x,t_1)\,\mathrm{d}x \right| \leqslant \varepsilon,$$

ce qui exprime une intégrabilité au sens de Riemann.

Nous pouvons maintenant rassembler toutes ces inégalités et achever la preuve du théorème de Fubini :

$$\begin{aligned}
&\left| \int_0^1 \left(\int_0^1 f(x,t)\,\mathrm{d}x \right) \mathrm{d}t - \sum_{i,j} |\mathrm{P}_{i,j}| f(t_{i,j}) \right| \\
&\qquad \leqslant \left| \int_0^1 \left(\int_0^1 f(x,t)\,\mathrm{d}x \right) \mathrm{d}t - \sum_{i=1}^{k} |\mathrm{Q}_i| \int_0^1 f(x,t_i)\,\mathrm{d}x \right| \\
&\qquad\qquad + \left| \sum_{i=1}^{k} |\mathrm{Q}_i| \int_0^1 f(x,t_i)\,\mathrm{d}x - \sum_{i,j} |\mathrm{P}_{i,j}| f(t_{i,j}) \right| \\
&\qquad \leqslant \varepsilon + \left| \sum_{i=1}^{k} \left(|\mathrm{Q}_i| \left[\int_0^1 f(x,t_i)\,\mathrm{d}x - \sum_{j=1}^{n_i} |\mathrm{P}_{i,j}| |\mathrm{Q}_i|^{-1} f(t_{i,j}) \right] \right) \right| \\
&\qquad \leqslant \varepsilon + \sum_{i=1}^{k} |\mathrm{Q}_i| \varepsilon = 2\varepsilon.
\end{aligned}$$

Cela montre que f est intégrable sur P et d'intégrale donnée par la formule de Fubini. □

La démonstration de la formule de Fubini pour les fonctions continues sur un pavé de dimension quelconque peut se faire par récurrence sur ladite dimension, en utilisant la formule démontrée en dimension 2. Nous laissons au lecteur le soin de rédiger cette récurrence. Enfin, le passage des fonctions continues aux fonctions Lebesgue-intégrables est relativement aisé si l'on utilise le théorème de densité des fonctions continues dans l'espace $\mathcal{L}^1$. Nous en donnerons l'argument dans le prochain chapitre.

Montrons sur un exemple comment utiliser la formule de Fubini. Imaginons que l'on veuille intégrer la fonction $f(x,y) = xy$ sur le triangle $\mathrm{T} = \{0 < x < 1, 0 < y < x\}$. Commençons par remarquer que f est continue sur T, que T est un ouvert borné, et que f se prolonge par continuité au bord de T. D'après le théorème 10.1, la fonction f est Lebesgue-intégrable sur T. Le triangle n'étant pas un pavé, il semble illusoire

de vouloir appliquer la formule de Fubini dans ce cas. Cependant, si l'on prolonge f par la valeur 0 hors de T, on obtient une fonction $\bar{f}^{\mathrm{T}}$ Lebesgue intégrable sur $\mathbf{R}^2$ à laquelle on peut appliquer le théorème de Fubini sur le carré unité $\mathrm{C} = [0,1] \times [0,1]$:

$$\begin{aligned}\int_{\mathrm{T}} xy \,\mathrm{d}x\,\mathrm{d}y &= \int_{\mathrm{C}} \bar{f}^{\mathrm{T}}(x,y)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_0^1 \left(\int_0^1 f(x,y)\,\mathrm{d}y \right) \mathrm{d}x \\ &= \int_0^1 \left(\int_0^x xy\,\mathrm{d}y \right) \mathrm{d}x \\ &= \int_0^1 x[y^2/2]_0^x \,\mathrm{d}x \\ &= \int_0^1 x^3/2 \,\mathrm{d}x \\ &= 1/8.\end{aligned}$$

Ainsi, lorsque l'on connaît la frontière de l'ouvert sur lequel on intègre, la formule de Fubini permet de calculer l'intégrale. Nous reverrons un peu plus loin, dans la démonstration de la formule de Green-Riemann, un autre exemple de cette méthode.

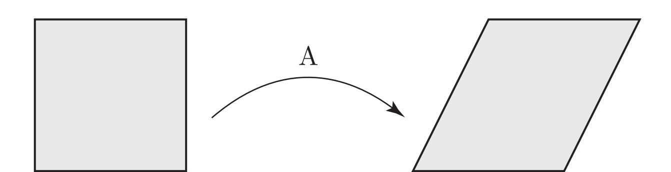
10.5. La formule de changement de variable

La formule de changement de variable est d'une grande importance en intégration. Elle permet par exemple de simplifier un domaine d'intégration pour faciliter la preuve d'estimées sinon inaccessibles ; elle est la matérialisation d'un lien profond entre la théorie de l'intégration et la géométrie de $\mathbf{R}^d$.

En une variable, nous avons facilement obtenu cette formule comme corollaire de la formule fondamentale. Nous aurions tout aussi bien pu la démontrer directement avec des subdivisions et des jauges, en exploitant le fait remarquable suivant : si $g : [a,b] \longrightarrow [a',b']$ est un difféomorphisme, alors g envoie une subdivision pointée de $[a,b]$ sur une subdivision pointée de $[a',b']$; cela tient au fait que tout difféomorphisme entre intervalles de $\mathbf{R}$ est soit croissant soit décroissant et que les pavés, c'est-à-dire, ici, les intervalles, sont définis par la relation d'ordre.

En dimension plus grande cela ne fonctionne plus du tout. La famille des pavés que nous avons choisie pour bâtir notre intégrale est basée sur un choix de coordonnées qui, dès qu'on applique un difféomorphisme, même simple, est chamboulé. Ainsi le pavé $\mathrm{P} = [0,1] \times [0,1]$ de $\mathbf{R}^2$ est envoyé sur un parallélogramme penché (qui n'est pas un pavé) par le difféomorphisme linéaire de matrice

$$\mathrm{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$



En d'autres termes, si $\mathcal{S} = \{(\mathrm{P}_1, t_1), \ldots, (\mathrm{P}_n, t_n)\}$ est une subdivision pointée du pavé P de $\mathbf{R}^d$ (avec $d \geqslant 2$) et si g est un difféomorphisme défini sur P alors d'une part $g(\mathrm{P})$ n'est pas nécessairement un pavé et, même si par hasard $g(\mathrm{P})$ en était un, la subdivision image $g(\mathcal{S}) = \{(g(\mathrm{P}_1), g(t_1)), \ldots, (g(\mathrm{P}_n), g(t_n))\}$ n'aurait quasiment aucune chance d'être une subdivision pointée.

Cette irruption de la géométrie dans le domaine de l'intégration va nous amener à donner de la formule de changement de variable une version plus faible que celle énoncée en une seule : nous devrons nous limiter au cas des fonctions Lebesgue-intégrables (et au cas des changements de variable donnés par des difféomorphismes) ; nous donnerons en effet un exemple de fonction f intégrable (mais par Lebesgue-intégrable) sur $[0,1] \times [0,1]$ telle que $f \circ \mathrm{A}$ n'est pas intégrable. En quelque sorte, l'action de la matrice A réarrange l'ordre de sommation de manière suffisamment violente pour faire diverger une série semi-convergente. Ce phénomène était absent en dimension 1 du fait de la préservation de l'ordre par difféomorphisme.

10.5.1. Prélude : le cas linéaire

Avant même d'énoncer la formule du changement de variable, examinons un cas simple qui nous guidera pour trouver l'énoncé. Soit $\mathrm{P} = [0,1] \times [0,1]$ le pavé unité de $\mathbf{R}^2$. On a alors

$$\int_{\mathrm{P}} \mathrm{d}x\, \mathrm{d}y = \int_{\mathbf{R}^2} \mathbf{1}_{\mathrm{P}}\, \mathrm{d}x\, \mathrm{d}y = \mathrm{Aire}(\mathrm{P}) = 1,$$

où $\mathbf{1}_{\mathrm{P}}$ est la fonction indicatrice de P. Soit maintenant A une matrice 2×2 à coefficients réels et inversible :

$$\mathrm{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2, \mathbf{R}).$$

Examinons la fonction $\mathbf{1}_{\mathrm{P}} \circ \mathrm{A}$: $\mathbf{1}_{\mathrm{P}} \circ \mathrm{A}(x,y) = 1$ si et seulement si $\mathrm{A}(x,y) \in \mathrm{P}$, soit donc $(x,y) \in \mathrm{A}^{-1}(\mathrm{P})$. On a donc la formule

$$\mathbf{1}_{\mathrm{P}} \circ \mathrm{A} = \mathbf{1}_{\mathrm{A}^{-1}(\mathrm{P})},$$

et $\mathrm{A}^{-1}(\mathrm{P})$ est un parallélogramme ; si l'on intègre $\mathbf{1}_{\mathrm{P}} \circ \mathrm{A}$ sur $\mathbf{R}^2$, nous trouvons alors

$$\begin{aligned} \int_{\mathbf{R}^2} \mathbf{1}_{\mathrm{P}} \circ \mathrm{A}\, \mathrm{d}x\, \mathrm{d}y &= \int_{\mathbf{R}^2} \mathbf{1}_{\mathrm{A}^{-1}(\mathrm{P})}\, \mathrm{d}x\, \mathrm{d}y \\ &= \int_{\mathrm{A}^{-1}(\mathrm{P})} \mathrm{d}x\, \mathrm{d}y \\ &= \mathrm{Aire}(\mathrm{A}^{-1}(\mathrm{P})). \end{aligned}$$

Mais l'on sait (*voir* un cours d'algèbre linéaire) que l'aire de $\mathrm{A}^{-1}(\mathrm{P})$ est égale à la valeur absolue du déterminant de A^{-1}. Nous avons donc démontré, dans un cas très simple, la formule de changement de variable :

$$\int_{\mathrm{P}} \mathbf{1}_{\mathrm{P}}(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathrm{A}^{-1}(\mathrm{P})} \mathbf{1}_{\mathrm{P}} \circ \mathrm{A}(u, v)|\det \mathrm{A}|\,\mathrm{d}u\,\mathrm{d}v,$$

ce qui n'est qu'une manière très compliquée d'écrire que $\mathrm{Aire}(\mathrm{A}^{-1}(\mathrm{P})) = |\det \mathrm{A}|^{-1}\mathrm{Aire}(\mathrm{P})$.

10.5.2. Fugue : le cas général

Dans ce qui précède et si l'on considère A comme une application différentiable de $\mathbf{R}^2$ dans lui-même, le déterminant de A n'est autre que le jacobien de l'application différentiable A.

Définition 10.9 — *Soient* Ω *un ouvert de* $\mathbf{R}^d$ *et* $\varphi : \Omega \longrightarrow \mathbf{R}^d$ *une application différentiable. Le jacobien de* φ *en un point* $\mathbf{x}$ *de* Ω *est le déterminant de la différentielle de* φ *au point* $\mathbf{x}$ *:*

$$(\mathrm{Jac}\,\varphi)(\mathbf{x}) = \det \mathrm{D}_{\mathbf{x}}\varphi = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1}(\mathbf{x}) & \frac{\partial \varphi_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial \varphi_1}{\partial x_d}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial \varphi_d}{\partial x_1}(\mathbf{x}) & \frac{\partial \varphi_d}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial \varphi_d}{\partial x_d}(\mathbf{x}) \end{vmatrix},$$

où les φ_i *sont les composantes de* φ.

Si φ est de classe $\mathcal{C}^1$, alors l'application $\mathrm{Jac}\,\varphi : \Omega \longrightarrow \mathbf{R}$ est continue, et en particulier localement intégrable.

Nous pouvons maintenant inférer la forme générale de la formule de changement de variable : soient Ω un ouvert de $\mathbf{R}^d$, $\varphi : \Omega \longrightarrow \varphi(\Omega) \subset \mathbf{R}^d$ un difféomorphisme de classe $\mathcal{C}^1$ et f une fonction Lebesgue-intégrable sur $\varphi(\Omega)$. Localement au voisinage d'un point $\mathbf{x}$ de Ω, φ se comporte comme l'application affine $v \mapsto \varphi(\mathbf{x}) + \mathrm{D}_{\mathbf{x}}\varphi(v)$. Elle transforme donc les volumes des petits pavés autour de $\mathbf{x}$ en les multipliant par $|(\mathrm{Jac}\,\varphi)(\mathbf{x})|$ et un passage à la limite permet alors de démontrer le théorème suivant :

Théorème 10.5 — *Soient* Ω *un ouvert de* $\mathbf{R}^d$, $\varphi : \Omega \longrightarrow \varphi(\Omega) \subset \mathbf{R}^d$ *un difféomorphisme de classe* $\mathcal{C}^1$ *et* $f : \varphi(\Omega) \longrightarrow \mathbf{R}$ *une fonction Lebesgue-intégrable. Alors* $f \circ \varphi$ *est Lebesgue-intégrable sur* Ω *et on a la formule*

$$\int_{\varphi(\Omega)} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\Omega} f \circ \varphi(\mathbf{u})|(\mathrm{Jac}\,\varphi)(\mathbf{u})|\,\mathrm{d}\mathbf{u}.$$

La mise en place du schéma de preuve évoqué ci-dessus est possible, mais assez longue et délicate. On peut, en se restreignant au cas où f est une fonction continue sur $\overline{\Omega}$

que l'on suppose aussi compact, donner une preuve un peu plus simple utilisant la formule de changement de variable pour les fonctions d'une seule variable ; on applique pour cela la formule de Fubini et une récurrence sur le nombre de variables affectées par le difféomorphisme φ. Malgré tout la démonstration reste fastidieuse, et nous la réservons à l'appendice.

Notons pour clore cette section que la formule de changement de variable permet d'exprimer une intégrale en coordonnées polaires, cylindriques ou sphériques, suivant le type de symétrie que présente le problème étudié.

Exemple : considérons un solide dont nous aimerions calculer le centre d'inertie. Un solide est la donnée d'un compact $\mathrm{K} \subset \mathbf{R}^3$ et d'une fonction continue $\rho : \mathrm{K} \longrightarrow]0, \infty[$ qui représente sa densité de masse, et que l'on étend à $\mathbf{R}^d$ en lui donnant la valeur 0 hors de K. Un tout petit cube $[x, x + \delta x] \times [y, y + \delta y] \times [z, z + \delta z]$ du solide a pour masse approximative $\rho(x, y, z)\delta x\, \delta y\, \delta z$. Ainsi la masse totale du solide est-elle donnée par l'intégrale triple :

$$\mathrm{M_K} = \iiint_{\mathbf{R}^3} \rho(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Le centre d'inertie $\boldsymbol{\xi}_\mathrm{K}$ du solide K est alors la moyenne du vecteur $\mathbf{x}$ relativement à la densité ρ ; en termes vectoriels nous écrirons

$$\boldsymbol{\xi}_\mathrm{K} = \frac{1}{\mathrm{M_K}} \iiint_{\mathbf{R}^3} \rho(\mathbf{x})\mathbf{x}\,\mathrm{d}\mathbf{x},$$

ou, si l'on note (ξ_1, ξ_2, ξ_3) les coordonnées de $\boldsymbol{\xi}_\mathrm{K}$:

$$\xi_1 = \frac{1}{\mathrm{M_K}} \iiint_{\mathbf{R}^3} x_1 \rho(x_1, x_2, x_3)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3,$$

avec une formule analogue pour ξ_2 et ξ_3. Si le solide (K, ρ) présente une invariance par symétrie, c'est-à-dire s'il existe une isométrie affine φ de $\mathbf{R}^3$ telle que $\varphi(\mathrm{K}) = \mathrm{K}$ et $\rho \circ \varphi = \rho$, alors le centre d'inertie de K doit être un point fixe de φ, c'est-à-dire vérifier $\varphi(\boldsymbol{\xi}_\mathrm{K}) = \boldsymbol{\xi}_\mathrm{K}$.

Prenons l'exemple simple d'un solide qui est une boule B de centre O et dont la densité de masse est radiale, c'est-à-dire ne dépend que de la distance au centre. Alors le solide est invariant par toutes les rotations de centre O et le centre d'inertie doit donc se situer en O. Un cas un peu plus compliqué est celui d'une demi-boule. Soit

$$\mathrm{D} = \left\{(x, y, z) \in \mathbf{R}^3, x^2 + y^2 + z^2 \leqslant \mathrm{R}^2 \text{ et } z \geqslant 0\right\}$$

dont on suppose la densité constante, égale à 1. Le solide D est invariant par toutes les rotations autour de l'axe des z et son centre d'inertie doit donc s'y situer. Autrement dit il nous suffit de calculer

$$\xi_z = \frac{1}{\mathrm{M_D}} \iiint_\mathrm{D} z\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

pour avoir la position de $\boldsymbol{\xi}_\mathrm{D}$. On commence par calculer la masse de la demi-boule. Celle-ci est de densité égale à 1, sa masse est donc égale au volume de la demi-boule. Mais la boule de rayon R a pour volume $4\pi\mathrm{R}^3/3$, et l'on a donc $\mathrm{M_D} = 2\pi\mathrm{R}^3/3$. On peut retrouver ce résultat en appliquant la formule de Fubini consistant à décomposer (x, y, z) en (x, y) et (z) dans le calcul du volume de la boule $\mathrm{B}(0, \mathrm{R})$:

$$\iiint_{\mathrm{B}(0,\mathrm{R})} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{-\mathrm{R}}^{+\mathrm{R}} \left(\iint_{\{x^2+y^2\leqslant \mathrm{R}^2-z^2\}} \mathrm{d}x\,\mathrm{d}y \right) \mathrm{d}z.$$

L'intégrale double n'est autre que l'aire du disque de rayon $\sqrt{\mathrm{R}^2 - z^2}$ et l'on a donc

$$\iiint_{\mathrm{B}(0,\mathrm{R})} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{-\mathrm{R}}^{+\mathrm{R}} \pi(\mathrm{R}^2 - z^2)\,\mathrm{d}z = \frac{4\pi\mathrm{R}^3}{3}.$$

La même méthode peut être appliquée pour calculer la cote du centre d'inertie de la demi-boule D :

$$\iiint_{\mathrm{D}} z\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^{\mathrm{R}} z\pi(\mathrm{R}^2 - z^2)\,\mathrm{d}z = \frac{\pi\mathrm{R}^4}{4},$$

soit finalement

$$\xi_z = \frac{3}{8}\mathrm{R},$$

une bien jolie formule. Mais on aurait pu exploiter autrement l'invariance du solide par rotation autour de l'axe des z, et utiliser les coordonnées cylindriques. De quoi s'agit-il ? Un point de coordonnées cartésiennes (x, y, z) de l'espace $\mathbf{R}^3$ est repéré, dans ce système, par sa cote z et, dans le plan déterminé par cette cote, par des coordonnées polaires (r, θ) de telle sorte que $x = r\cos\theta$ et $y = r\sin\theta$. Un solide de révolution est alors simplement donné par une fonction continue positive $f : [a, b] \longrightarrow \mathbf{R}_+$, ce qui permet de définit le compact $\mathrm{K}_f = \{(r, \theta, z),\ 0 \leqslant r \leqslant f(z),\ a \leqslant z \leqslant b\}$, et d'une densité $\rho(r, z)$ définie pour tout (r, z) tel que $a \leqslant z \leqslant b$ et $0 \leqslant r \leqslant f(z)$. Le jacobien de la transformation permettant le passage des coordonnées cartésiennes aux coordonnées polaire vaut r, et l'on peut donc calculer la masse du solide de révolution par la formule :

$$\mathrm{M_K} = \iiint_{\mathrm{K}} \rho(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_a^b \int_0^{f(z)} \int_0^{2\pi} r\rho(r, z)\,\mathrm{d}\theta\,\mathrm{d}r\,\mathrm{d}z.$$

Si la densité est constante égale à 1, on trouve pour masse

$$\mathrm{M_K} = \pi \int_a^b f(z)^2\,\mathrm{d}z$$

et l'on peut écrire une formule analogue pour la cote du centre d'inertie.

Pour conclure cette partie, donnons la forme générale des *changements de variable classiques*. Dans le plan on peut repérer les points par leurs coordonnées cartésiennes (x, y) ou par leurs coordonnées polaires (r, θ), avec $r > 0$ et $\theta \in [0, 2\pi[$, de sorte

que $(x, y) = (r\cos\theta, r\sin\theta)$. Si f est une fonction Lebesgue-intégrable sur le disque $\mathrm{D}(0, r_0)$ centré en 0 et de rayon r_0 alors on peut effectuer un changement de variable et exprimer f en coordonnées polaires. Nous noterons

$$f(x, y) = f(r\cos\theta, r\sin\theta) = f_p(r, \theta),$$

de sorte que f_p est définie sur le pavé $]0, r_0[\times[0, 2\pi[$. La formule de passage aux coordonnées polaires pour les intégrales doubles s'écrit alors :

$$\int_{\mathrm{D}(0,r_0)} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{r_0}\int_0^{2\pi} f_p(r, \theta) r\,\mathrm{d}r\,\mathrm{d}\theta.$$

Dans l'espace $\mathbf{R}^3$, on dispose de deux systèmes de coordonnées en sus des coordonnées cartésiennes (x, y, z) : les coordonnées cylindriques et les coordonnées sphériques. Le point (x, y, z) est donné en coordonnées cylindriques par (r, θ, z), où (r, θ) sont les coordonnées polaires du point du plan (x, y), de sorte que

$$(x, y, z) = (r\cos\theta, r\sin\theta, z).$$

Ses coordonnées sphériques sont données par (r, θ, φ), où (r, θ) sont encore les coordonnées polaires du point (x, y) mais où, cette fois, (r, φ) sont en quelque sorte les coordonnées polaires du point (x, y, z) dans le plan vertical contenant l'axe des z et le point $(x, y, 0)$. De cette manière l'on a

$$(x, y, z) = (r\cos\theta\sin\varphi, r\sin\theta\sin\varphi, r\cos\varphi).$$

Si f est définie sur une partie de $\mathbf{R}^3$, nous noterons

$$f(x, y, z) = f(r\cos\theta, r\sin\theta, z) = f_c(r, \theta, z)$$

et

$$f(x, y, z) = f(r\cos\theta\sin\varphi, r\sin\theta\sin\varphi, r\cos\varphi) = f_s(r, \theta, \varphi).$$

Les coordonnées cylindriques sont bien adaptées aux fonctions définies sur un *domaine de révolution*. Un tel domaine Ω_c est défini par les inégalités

$$\theta \in [0, 2\pi[,\ z_1 \leqslant z \leqslant z_2 \text{ et } 0 \leqslant r \leqslant c(z),$$

où c est une fonction continue positive définie sur l'intervalle $[z_1, z_2]$. La formule de passage aux coordonnées cylindriques pour les intégrales triples s'écrit alors, pour toute fonction f Lebesgue-intégrable sur Ω_c :

$$\int_{\Omega_c} f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{z_1}^{z_2}\int_0^{c(z)}\int_0^{2\pi} f_c(r, \theta, z) r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}z.$$

Les coordonnées sphériques sont bien adaptées aux fonctions définies sur des ouverts en coquille Ω_{r_1,r_2} définis par les conditions

$$r_1 \leqslant r \leqslant r_2,\ \theta \in [0, 2\pi[\ \text{ et }\ \varphi \in [0, \pi[.$$

Si f est Lebesgue-intégrable sur Ω_{r_1,r_2}, alors la formule de passage aux coordonnées sphériques pour les intégrales triples s'écrit :

$$\int_{\Omega_{r_1,r_2}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{r_1}^{r_2}\int_0^{2\pi}\int_0^{\pi} f_s(r,\theta,\varphi)r^2\sin\varphi\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi.$$

On peut bien sûr étendre toutes ces formules à des domaines définis dans divers secteurs angulaires et nous renvoyons aux exercices pour quelques applications pratiques.

10.6. La formule de Green-Riemann

La formule fondamentale du calcul différentiel et intégral offre à l'analyste un outil formidable qui fait le pont entre l'intégration et la dérivation. Il fournit ainsi à l'étude de sujets comme les équations différentielles tout l'arsenal puissant de la théorie de l'intégration. Il existe un analogue de cette formule en plusieurs variables, c'est la *formule de Stokes.* Ne serait-ce qu'énoncer celle-ci dans le cas général nous mènerait trop loin, dans le territoire des variétés et des formes différentielles. Nous allons cependant ici donner un avatar de cette formule valable pour certains domaines simples du plan $\mathbf{R}^2$.

Dans la suite, nous identifierons $\mathbf{R}^2$ à la droite complexe $\mathbf{C}$, et utiliserons de manière tacite cette identification. Ainsi, si (x_0, y_0) est un point de $\mathbf{R}^2$, nous le noterons $z_0 = x_0 + iy_0$ comme nombre complexe, et utiliserons librement les notions de module et d'argument.

Avant même d'énoncer la formule de Green-Riemann, il nous faut passer par quelques préliminaires de géométrie des arcs dans le plan, afin de pouvoir définir l'intégration sur iceux.

10.6.1. Chemins

Définition 10.10 — *Un chemin (ou arc) de classe $\mathcal{C}^1$ du plan $\mathbf{R}^2$ est une application γ, de classe $\mathcal{C}^1$, définie sur un intervalle compact $[a,b]$ de $\mathbf{R}$ et à valeurs dans $\mathbf{R}^2$, et* de dérivée partout non nulle. *Le chemin γ est dit continu $\mathcal{C}^1$ par morceaux si γ est continue, et est $\mathcal{C}^1$ par morceaux au sens qui précède : c'est un recollement bout à bout d'un nombre fini de chemins de classe $\mathcal{C}^1$.*

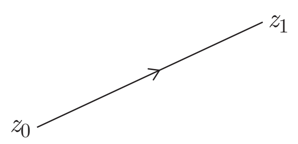
Exemples :

– Soient (x_0, y_0) et (x_1, y_1) deux points du plan. Le segment les reliant est un chemin $\mathcal{C}^1$, qu'on peut paramétrer par une combinaison barycentrique :

$$\gamma(t) = ((1-t)x_0 + tx_1, (1-t)y_0 + ty_1),\ \ t \in [0,1].$$

En notation complexe, avec $z_0 = x_0 + iy_0, z_1 = x_1 + iy_1$, cela s'écrit

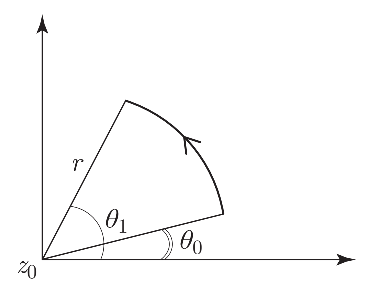
$$\gamma(t) = (1-t)z_0 + tz_1.$$



Ainsi, on parcourt le segment $[z_0, z_1]$ en une unité de temps, c'est à dire à vitesse constante $|z_1 - z_0|$.

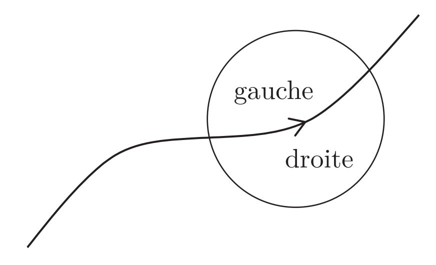
– Soient $0 \leqslant \theta_0 < \theta_1 \leqslant 2\pi$ et $\gamma : [\theta_0, \theta_1] \longrightarrow \mathbf{C}$ définie par

$$\gamma(t) = z_0 + r\mathrm{e}^{it}.$$



C'est un chemin $\mathcal{C}^1$, qui parcourt l'arc de cercle centré en z_0, de rayon r, et d'angles compris entre θ_0 et θ_1, dans le sens trigonométrique.

Remarque : le paramétrage oriente le chemin. En particulier, tout chemin continu $\mathcal{C}^1$ par morceaux, admet au voisinage de chacun des points où sa vitesse est non nulle un côté gauche et un côté droit.



– Attention, la condition de non-nullité de la dérivée est importante pour éviter des singularités topologiques. Ainsi, la courbe $(t^3 \sin(1/t), t^3 \cos(1/t))$ est-elle de classe $\mathcal{C}^1$, mais de dérivée nulle en zéro : elle y « spirale », rendant impossible de définir, par exemple, la notion de droite ou de gauche de la courbe au voisinage de l'origine.

– Soit n un entier naturel non nul. Le chemin $\mathcal{C}^1$ défini par

$$\gamma(t) = \mathrm{e}^{int}, \ \ t \in [0, 2\pi],$$

représente le cercle unité parcouru n fois dans le sens positif. Remarquons que le même objet peut être représenté par

$$\eta(t) = \mathrm{e}^{it}, \ \ t \in [0, 2n\pi].$$

Définition 10.11 — *Deux chemins $\gamma_1 : [a, b] \longrightarrow \mathbf{R}^2$ et $\gamma_2 : [a', b'] \longrightarrow \mathbf{R}^2$ sont dits équivalents s'il existe un difféomorphisme de classe $\mathcal{C}^1$ et croissant $\varphi : [a', b'] \longrightarrow [a, b]$ (c.-à-d. un changement de paramètre) tel que*

$$\forall t \in [a', b'], \ \ \gamma_2(t) = \gamma_1(\varphi(t)).$$

Ainsi, dans l'exemple du cercle parcouru n fois décrit ci-dessus, les deux chemins γ et η sont équivalents. De même, si l'on paramètre l'arc de cercle centré en z_0 et de rayon r vu plus haut par

$$\gamma_2(t) = z_0 + r\mathrm{e}^{2i\pi t}, \ \ t \in [\theta_0/2\pi, \theta_1/2\pi],$$

on obtient un chemin équivalent où, au lieu de compter en radians, on compte en nombre de tours.

Moralité : dans la suite, quand on dira « soit γ un arc continu, $\mathcal{C}^1$ par morceaux », on entendra toujours l'objet géométrique, en sous-entendant un paramétrage. Pour être plus précis, un chemin sera en fait une classe d'équivalence de paramétrage. Pour ne pas alourdir les notations, nous noterons encore γ l'image de l'application $\gamma : [a, b] \longrightarrow \mathbf{R}$, et ainsi $z \in \gamma$ signifiera simplement que $z \in \gamma([a, b])$.

On pourra de la sorte parler du segment $[z_0, z_1]$, ou encore du « cercle unité parcouru n fois dans le sens positif », et un dessin pourra suffire à le décrire.

Dans ce dernier cas, il y a encore une ambiguïté dont il faut se débarrasser : le choix d'un point de départ sur le cercle.

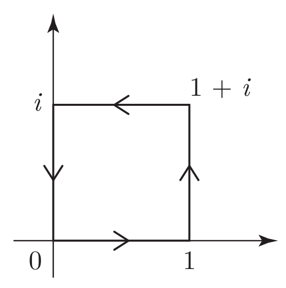
Définition 10.12 — *On appelle lacet (ou chemin fermé, ou arc fermé...) tout chemin continu, de classe $\mathcal{C}^1$ par morceaux, dont l'origine est égal au but. Plus précisément, si $\gamma : [a, b] \longrightarrow \mathbf{R}^2$ est un chemin continu de classe $\mathcal{C}^1$ par morceaux, on dit que c'est un lacet si $\gamma(a) = \gamma(b)$. Le lacet γ est dit simple s'il ne se recoupe pas, en d'autres termes si γ, restreint à $]a, b[$, est injectif.*

Soit γ un lacet. Quitte à effectuer un changement affine de paramètre, on peut supposer que γ est défini sur $[0, 1]$ et que $\gamma(0) = \gamma(1)$. On peut enfin prolonger γ périodiquement à $\mathbf{R}$ en entier, et on obtient une application $\bar{\gamma} : \mathbf{R} \longrightarrow \mathbf{R}^2$ qui est continue, $\mathcal{C}^1$ par morceaux et 1-périodique. Cela permet de définir un paramétrage

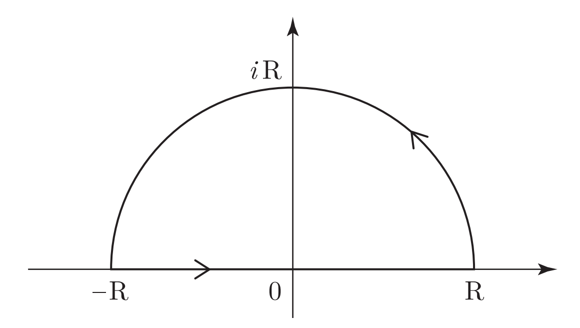
$\widehat{\gamma} : \mathbf{R}/\mathbf{Z} \longrightarrow \mathbf{R}^2$, qui est aussi continu et $\mathcal{C}^1$ par morceaux. Tout lacet est ainsi donné par une application continue et $\mathcal{C}^1$ par morceaux de $\mathbf{R}/\mathbf{Z}$ dans $\mathbf{R}^2$. Deux tels lacets seront dits équivalents si, après un éventuel changement de paramètre donné par un difféomorphisme positif de $\mathbf{R}/\mathbf{Z}$, les deux applications coïncident à une translation du paramètre près. Le lecteur assidu tentera d'écrire les détails formels.

Exemples :

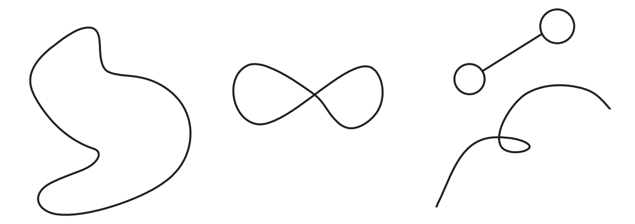
– Le carré unité, parcouru une fois dans le sens positif, est un lacet continu et $\mathcal{C}^1$ par morceaux.



– Voici un autre exemple, constitué d'un grand segment $[-\mathrm{R}, +\mathrm{R}]$ inclus dans l'axe réel, et du demi-cercle centré en 0 et de rayon R allant de $-\mathrm{R}$ à $+\mathrm{R}$, que l'on retrouve dans certains calculs d'intégrales par la méthode des résidus issue de l'analyse complexe.



– Quelques exemples et contre-exemples de lacets simples.



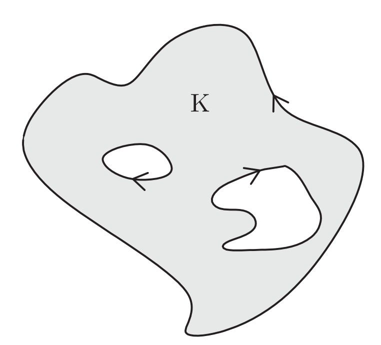
10.6.2. Domaines

La formule de Green-Riemann est une égalité entre une intégrale sur un lacet (nous verrons plus loin de quoi il s'agit) et une intégrale sur le domaine bordé par ce lacet. Plus généralement, nous considérerons des compacts réguliers du plan :

Définition 10.13 — *Un compact* $\mathrm{K} \subset \mathbf{C}$ *est dit régulier si son bord est l'union d'un nombre fini de lacets simples, continus et* $\mathcal{C}^1$ *par morceaux, disjoints deux à deux :*

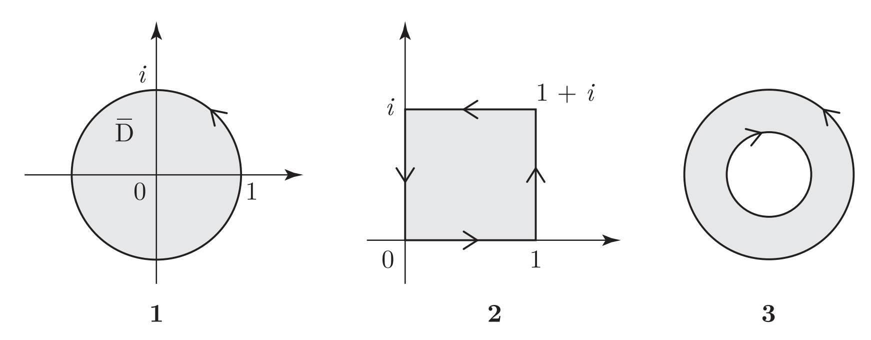
$$\partial \mathrm{K} = \gamma_1 \cup \cdots \cup \gamma_n.$$

Par convention, on oriente chacune des composantes γ_i *du bord pour que* K *soit, en tout point de* γ_i, *« à gauche » de* γ_i.



Exemples :

1. Le disque unité fermé $\overline{\mathrm{D}}$ est un compact régulier, dont le bord n'a qu'une composante, le cercle unité parcouru une fois dans le sens trigonométrique.
2. Le carreau unité fermé est un compact régulier, dont le bord est le carré unité parcouru une fois dans le sens positif.
3. La couronne fermée $\mathrm{C} = \{z \in \mathbf{C}, 1 \leqslant |z| \leqslant 2\}$ est un compact régulier, dont le bord a deux composantes : le cercle de centre 0 de rayon 2, parcouru dans le sens trigonométrique, et le cercle unité, parcouru dans le sens chronométrique.



10.6.3. 1-formes différentielles

Nous ne rentrons pas ici dans les détails de la définition générale des formes différentielles, mais les décrivons plutôt comme des objets algébriques sujets à certaines opérations naturelles, en particulier l'intégration sur des chemins. Soient U un ouvert de $\mathbf{R}^2$ et f une fonction, disons de classe $\mathcal{C}^2$, définie sur U. La différentielle de f au point (x, y) de U est, par définition, la forme linéaire sur $\mathbf{R}^2$ donnée par

$$\mathrm{D}f(x,y) = \frac{\partial f}{\partial x}(x,y)\mathrm{e}_x^* + \frac{\partial f}{\partial y}(x,y)\mathrm{e}_y^*$$

où e_x^* et e_y^* sont les deux formes de coordonnées définies par

$$\mathrm{e}_x^*\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = a \quad \text{et} \quad \mathrm{e}_y^*\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = b.$$

Il est d'usage en calcul différentiel de noter ces deux formes linéaires $\mathrm{e}_x^* = \mathrm{d}x$ et $\mathrm{e}_y^* = \mathrm{d}y$ de sorte que

$$\mathrm{D}f = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y.$$

La différentielle de f définit donc une application de classe $\mathcal{C}^1$ de U à valeurs dans le dual de $\mathbf{R}^2$: c'est un *champ de formes linéaires*. La définition suivante est alors naturelle :

Définition 10.14 — *Une* 1*-forme différentielle définie sur un ouvert* U *de* $\mathbf{R}^2$ *est une application de classe* $\mathcal{C}^1$ *de* U *à valeurs dans le dual de* $\mathbf{R}^2$*. Les formes* $\{\mathrm{d}x, \mathrm{d}y\}$ *formant une base du dual, une forme différentielle* α *est de la forme*

$$\alpha = \mathrm{P}\,\mathrm{d}x + \mathrm{Q}\,\mathrm{d}y,$$

où P, Q *sont deux fonctions de classe* $\mathbf{C}^1$*, définies sur* U *et à valeurs dans* $\mathbf{R}$ *ou* $\mathbf{C}$*.*

Voici quelques exemples :

$$\alpha = \mathrm{d}x + \mathrm{d}y,\ \mathrm{U} = \mathbf{R}^2,$$
$$\alpha = \mathrm{d}x + i\,\mathrm{d}y,\ \mathrm{U} = \mathbf{R}^2,$$

ou encore

$$\alpha = \frac{\mathrm{d}x}{x} - i\frac{\mathrm{d}y}{y},\ \mathrm{U} = \mathbf{R}^2 - \{xy = 0\}.$$

Remarquons tout de suite que l'on peut naturellement additionner deux 1-formes, les multiplier par une fonction de classe $\mathcal{C}^1$, et qu'ainsi l'espace $\Lambda^1(\mathrm{U})$ des 1-formes différentielles définies sur U est un $\mathbf{C}$-espace vectoriel (et même un $\mathcal{C}^1(\mathrm{U})$-module). Mais ce qui sera plus important pour nous, c'est que l'on peut intégrer les 1-formes le long des chemins.

10.6.4. Intégration sur un chemin

Définition 10.15 — *Soient* $\alpha = \mathrm{P}\,\mathrm{d}x + \mathrm{Q}\,\mathrm{d}y$ *une* 1*-forme différentielle définie sur un ouvert* U *de* **C***, et* $\gamma : [a, b] \longrightarrow \mathrm{U}$ *un chemin de classe* $\mathcal{C}^1$ *de* U*. On appelle intégrale de* α *le long de* γ *la quantité*

$$\int_\gamma \alpha = \int_\gamma \mathrm{P}\,\mathrm{d}x + \mathrm{Q}\,\mathrm{d}y = \int_a^b \big[\mathrm{P}(x(t), y(t))x'(t) + \mathrm{Q}(x(t), y(t))y'(t)\big]\,\mathrm{d}t,$$

où $\gamma(t) = (x(t), y(t))$*.*

Remarquons que cela revient à poser, le long de γ :

$$\mathrm{d}x = x'(t)\,\mathrm{d}t \quad \text{et} \quad \mathrm{d}y = y'(t)\,\mathrm{d}t.$$

On peut donner une interprétation physique à l'intégrale d'une 1-forme le long d'un chemin. Soit en effet, en conservant les notations ci-dessus, le champ de vecteurs défini par

$$\vec{\mathrm{A}}(x, y) = (\mathrm{P}(x, y), \mathrm{Q}(x, y)).$$

Alors

$$\int_\gamma \alpha = \int_a^b \left\langle \vec{\mathrm{A}}(\gamma(t)), \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right\rangle \mathrm{d}t,$$

et donc l'intégrale de α le long de γ représente le travail, ou circulation, du champ de forces $\vec{\mathrm{A}}$ le long du chemin γ.

Exemples :

1. Soit $\alpha = \mathrm{d}x$ et γ le segment entre 0 et 1. On a alors

$$\int_\gamma \alpha = \int_0^1 x'(t)\,\mathrm{d}t = \int_0^1 1\,\mathrm{d}t = 1.$$

2. Commençons par noter $\mathrm{d}z = \mathrm{d}(x + iy) = \mathrm{d}x + i\,\mathrm{d}y$. Soient alors $f = u + iv : \mathrm{U} \longrightarrow \mathbf{C}$ une fonction de classe $\mathcal{C}^1$, et $\gamma : [a, b] \longrightarrow \mathrm{U}$ un chemin $\mathcal{C}^1$ de U. On a

$$\begin{aligned} f(z)\,\mathrm{d}z &= (u(x, y) + iv(x, y))(\,\mathrm{d}x + i\,\mathrm{d}y) \\ &= (u(x, y) + iv(x, y))\,\mathrm{d}x + i(u(x, y) + iv(x, y))\,\mathrm{d}y, \end{aligned}$$

ce qui donne

$$\int_\gamma f(z)\,\mathrm{d}z = \int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}t,$$

où $\gamma'(t) = x'(t) + iy'(t)$. Ainsi, on a fait l'identification $\mathrm{d}z = \gamma'(t)\,\mathrm{d}t$ le long de l'arc γ.

3. Soit $\gamma(t) = \mathrm{e}^{it}, t \in [0, 2\pi]$ le cercle unité parcouru 1 fois dans le sens positif. Soient par ailleurs $n \in \mathbf{Z}$ et $f_n : \mathbf{C}^* \longrightarrow \mathbf{C}^*$ la fonction définie par

$$f_n(x) = z^n.$$

On voit aisément que $\gamma'(t) = i\mathrm{e}^{it}$, et donc que

$$\int_\gamma f_n(z)\,\mathrm{d}z = i\int_0^{2\pi} \mathrm{e}^{i(n+1)t}\,\mathrm{d}t,$$

qui est nulle si $n \neq -1$ et vaut $2i\pi$ si $n = -1$.

La première question qui vient à l'esprit après ces quelques exemples est celle de savoir si l'intégrale d'une forme le long d'une chemin γ dépend du paramétrage choisi pour ce chemin, ou si, plus intrinsèquement, elle ne dépend que de l'objet géométrique sous-jacent.

Théorème 10.6 — *Soit α est une 1-forme différentielle définie sur un ouvert U, et si γ_1 et γ_2 sont deux chemins de U équivalents, alors*

$$\int_{\gamma_1} \alpha = \int_{\gamma_2} \alpha.$$

Démonstration — Si $\gamma_i, i = 1, 2$ est défini sur $[a_i, b_i]$, on se ramène à des intégrales sur les segments $[a_i, b_i]$ et on applique la formule de changement de variable pour les intégrales de fonctions d'une variable. □

Corollaire 10.7 — *L'intégrale d'une 1-forme le long d'un chemin ne dépend pas du paramétrage de ce dernier.*

Quelques propriétés de l'intégrale de chemin

Remarquons tout d'abord que la définition de l'intégrale d'une 1-forme sur un chemin de classe $\mathcal{C}^1$ se généralise naturellement aux chemins continus, $\mathcal{C}^1$ par morceaux : on met bout à bout des intervalles $\mathcal{C}^1$, et le résultat de l'intégrale est donné par une relation de Chasles. Ce que nous avons dit jusqu'à maintenant, en particulier le théorème 10.6, est valable dans ce cadre. Cela permet aussi de définir l'intégrale d'une 1-forme sur le bord d'un compact régulier : si K est un compact régulier, inclus dans le domaine de définition d'une 1-forme différentielle f, et si $\partial \mathrm{K} = \gamma_1 \cup \cdots \gamma_n$, alors

$$\int_{\partial \mathrm{K}} \alpha = \sum_{i=1}^{n} \int_{\gamma_i} \alpha.$$

Enfin, l'application qui à une 1-forme définie sur U associe son intégrale le long d'un chemin de U donné est une application linéaire de $\Lambda^1(\mathrm{U})$ dans $\mathbf{R}$ ou $\mathbf{C}$.

Passons maintenant à une propriété utile par ailleurs. Un chemin $\gamma = x+iy : [a,b] \longrightarrow \mathbf{R}^2$ de classe $\mathcal{C}^1$ est rectifiable, et sa longueur est donnée par la formule

$$\begin{aligned}\operatorname{long}(\gamma) &= \int_a^b \left\| \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right\| \mathrm{d}t \\ &= \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\,\mathrm{d}t \\ &= \int_a^b |\gamma'(t)|\,\mathrm{d}t,\end{aligned}$$

où $\gamma'(t) = x'(t) + iy'(t)$. La longueur ne dépend bien sûr pas du paramétrage choisi. Cette formule admet une généralisation immédiate aux chemins continus, $\mathcal{C}^1$ par morceaux. Soit maintenant $\alpha = \mathrm{P}\,\mathrm{d}x + \mathrm{Q}\,\mathrm{d}y$ une 1-forme définie au voisinage de $\gamma([a,b])$. On a alors

$$\begin{aligned}\left| \int_\gamma \alpha \right| &= \left| \int_a^b \left\langle \mathrm{P}(\gamma(t))\mathrm{Q}(\gamma(t)), \frac{\mathrm{d}\gamma}{\mathrm{d}t} \right\rangle \mathrm{d}t \right| \\ &\leqslant \operatorname{long}(\gamma) \max \left\{ |\mathrm{P}(\gamma(t)) + i\mathrm{Q}(\gamma(t))|,\, t \in [a,b] \right\}.\end{aligned}$$

Dans le cas particulier où $\alpha = f(z)\,\mathrm{d}z$, avec f une fonction réelle ou complexe définie sur un voisinnage de $\gamma([a,b])$, cette inégalité s'écrit

$$\left| \int_\gamma f(z)\,\mathrm{d}z \right| \leqslant \max \left\{ |f(z)|, z \in \gamma \right\} \operatorname{long}(\gamma).$$

10.6.5. La formule de Green-Riemann

Démontrer cette formule en toute généralité est difficile et assez long, et c'est pourquoi nous allons plutôt nous limiter à un cas plus simple (où l'on fait une hypothèse de régularité sur la forme différentielle que l'on intègre) et où la topologie est cachée (en intégrant sur des bords de domaines).

Théorème 10.8 (Formule de Green-Riemann) — *Soient* K *un compact à bord régulier de* $\mathbf{R}^2$*,* $\alpha = \mathrm{P}\,\mathrm{d}x + \mathrm{Q}\,\mathrm{d}y$ *une* 1*-forme différentielle définie sur un voisinage de* K*. Alors*

$$\int_{\partial \mathrm{K}} (\mathrm{P}\,\mathrm{d}x + \mathrm{Q}\,\mathrm{d}y) = \iint_{\mathrm{K}} \left(\frac{\partial \mathrm{Q}}{\partial x} - \frac{\partial \mathrm{P}}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y.$$

Avant de démontrer cette formule, donnons-en un corollaire simple. On dit que la 1-forme différentielle α est exacte s'il existe une fonction φ (définie là où α l'est), de classe $\mathcal{C}^2$ et telle que

$$\alpha = \mathrm{d}\varphi = \frac{\partial \varphi}{\partial x}\,\mathrm{d}x + \frac{\partial \varphi}{\partial y}\,\mathrm{d}y.$$

Corollaire 10.9 — *Si α est une 1-forme différentielle exacte, alors*

$$\int_{\partial K} \alpha = 0.$$

Démonstration — On applique simplement la formule de Green-Riemann :

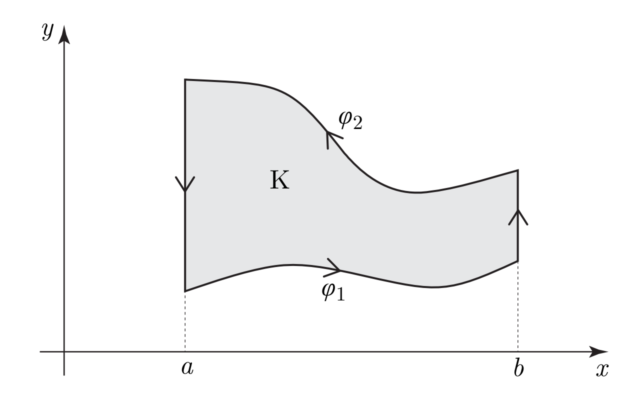
$$\int_{\partial K} \alpha = \int_{\partial K} \left(\frac{\partial \varphi}{\partial x} \, dx + \frac{\partial \varphi}{\partial y} \, dy \right) = \iint_K \left(\frac{\partial^2 \varphi}{\partial x \partial y} - \frac{\partial^2 \varphi}{\partial y \partial x} \right) dx \, dy.$$

L'intégrande dans la dernière intégrale est nulle d'après le théorème de Schwarz et parce que φ est de classe $\mathcal{C}^2$. □

Passons maintenant à la démonstration de la formule de Green-Riemann. On commence par la démontrer dans le cas particulier où le compact K est de la forme :

$$K = \left\{ (x, y), \ \varphi_1(x) \leqslant y \leqslant \varphi_2(x), \ x \in [a, b] \right\},$$

où φ_1 et φ_2 sont deux fonctions de classe $\mathcal{C}^1$ définies sur $[a, b]$ vérifiant $\varphi_1 < \varphi_2$.



Le compact K est un compact régulier dont le bord est un chemin continu, $\mathcal{C}^1$ par morceaux, que l'on peut décomposer en 4 parties :

- γ_1 le graphe de φ_1 parcouru dans le sens des x croissants. On peut le paramétrer par
$$\gamma_1(t) = (t, \varphi_1(t)), \ t \in [a, b] \,;$$
- γ_2 le segment entre $(b, \varphi_1(b))$ et $(b, \varphi_2(b))$;
- γ_3 le graphe de φ_2, parcouru dans le sens des x décroissants, que l'on peut paramétrer par
$$\gamma_3(t) = (-t, \varphi_2(-t)), \ t \in [-b, -a] \,;$$
- et enfin γ_4 le segment entre $(a, \varphi_2(a))$ et $(a, \varphi_1(a))$.

D'après la formule de Fubini, on a, comme P est de classe $\mathcal{C}^1$, que

$$\iint_K \frac{\partial P}{\partial y}\,dx\,dy = \int_a^b \left[\int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y}\,dy\right] dx$$
$$= \int_a^b \left[P(x,\varphi_2(x)) - P(x,\varphi_1(x))\right] dx$$
$$= \int_a^b P(x,\varphi_2(x))\,dx - \int_a^b P(x,\varphi_1(x))\,dx.$$

La deuxième égalité découle de la formule fondamentale du calcul différentiel et intégral. Par définition, cela donne

$$\iint_K \frac{\partial P}{\partial y}\,dx\,dy = -\int_{\gamma_1} P\,dx - \int_{\gamma_3} P\,dx.$$

Par ailleurs, le long de γ_2 et γ_4, qui sont verticaux, on a $dx = 0$, et l'on en déduit donc que

$$\iint_K \frac{\partial P}{\partial y}\,dx\,dy = -\int_{\partial K} P\,dx.$$

Si l'on veut une formule analogue pour $Q\,dy$, il nous faut faire tourner la figure d'un quart de tour. Supposons donc maintenant K' donné par

$$K' = \big\{(x,y),\ \psi_1(y) \leqslant x \leqslant \psi_2(y), y \in [c,d]\big\}.$$

Par le même calcul que précédemment, on montre que

$$\iint_{K'} \frac{\partial Q}{\partial x}\,dx\,dy = \int_{\partial K'} Q\,dy.$$

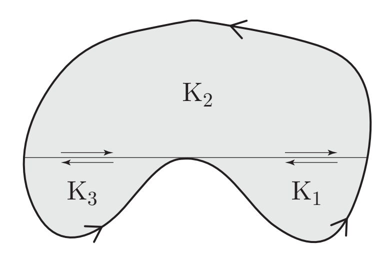
Supposons maintenant que K puisse être défini à la fois comme dans les deux cas ci-avant. Le disque unité est un exemple d'un tel compact. Comme les deux formules que l'on vient de démontrer pour un tel compact sont vraies, on en déduit aisément la formule de Green-Riemann dans ce cas :

$$\int_{\partial K} (P\,dx + Q\,dy) = \iint_K \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy.$$

La démonstration dans le cas général s'obtient en appliquant le lemme suivant.

Lemme 10.10 — *Tout compact régulier est union finie de compacts vérifiant à la fois les deux hypothèses plus haut, ayant des bords communs parallèles aux axes.*

Nous admettrons ce lemme, mais voici un dessin qui l'illustre :



Les intégrales sur les côtés communs se compensent, car ils sont parcourus une fois dans un sens puis une fois dans l'autre. Ainsi, on a

$$\int_{\partial K} P\,dx + Q\,dy = \int_{\partial K_1} P\,dx + Q\,dy + \int_{\partial K_2} P\,dx + Q\,dy + \int_{\partial K_3} P\,dx + Q\,dy.$$

la formule de Green-Riemann s'ensuit puisque par ailleurs, comme les K_i sont d'intérieur disjoints et leurs frontières communes sont de mesure de Lebesgue nulle dans le plan, on a

$$\iint_K = \iint_{K_1} + \iint_{K_2} + \iint_{K_3}.$$

Un exemple d'application : soit K un compact régulier. D'après la formule de Green-Riemann, on a

$$\int_{\partial K} x\,dy = \iint_K dx\,dy = -\int_{\partial K} y\,dx.$$

On en déduit que

$$\iint_K dx\,dy = \text{Aire}(K) = \frac{1}{2}\int_{\partial K}(x\,dy - y\,dx).$$

La formule de Green-Riemann a de nombreuses applications en mécanique des fluides, en électrostatique ou magnétostatique. Mais c'est en analyse complexe qu'elle trouve sa plus belle application : elle a pour corollaire le théorème central de ce domaine, la formule intégrale de Cauchy.

10.6.6. La formule de la divergence

Pour clore cette section, mentionnons une formule qui est analogue à celle de Green-Riemann mais valable en toute dimension. On peut la démontrer en s'appuyant sur la formule fondamentale du calcul différentiel et intégral grâce à une récurrence sur la dimension et à la formule de Fubini, comme nous l'avons fait pour la formule de Green. Nous ne donnons pas les définitions de manière complète, ce qui prendrait

du temps. Soit Ω un ouvert borné de $\mathbf{R}^d$ dont le bord $\partial\Omega$ est une hypersurface de classe $\mathcal{C}^\infty$. Cela signifie grossièrement qu'au voisinage de chaque point de $\partial\Omega$ il existe un système de coordonnées $(x_1, \ldots, x_d)$ tel que $\partial\Omega$ puisse être paramétré par une fonction f de classe $\mathcal{C}^\infty$ d'un ouvert U de $\mathbf{R}^{d-1}$ vers $\mathbf{R}$, de sorte que, dans U, $\partial\Omega$ soit la portion de nappe

$$\partial\Omega \cap (\mathrm{U} \times]-\varepsilon, \varepsilon[) = \{(x_1, \ldots, x_{d-1}, f(x_1, \ldots, x_{d-1})),\ (x_1, \ldots, x_{d-1}) \in \mathrm{U}\}.$$

On doit imposer des conditions techniques à la fonction f comme, par exemple, d'être de différentielle partout non nulle. Sous ces conditions, le bord de Ω est une hypersurface compacte sur laquelle on va pouvoir intégrer certains objets : les formes différentielles d'ordre $d-1$. Comme nous ne voulons pas nous aventurer sur ce terrain, nous allons plutôt définir le *flux* d'un champ de vecteurs au travers de $\partial\Omega$. Soit donc $\overrightarrow{\mathrm{X}}$ une application de classe $\mathcal{C}^1$ définie sur un ouvert contenant $\overline{\Omega}$ et à valeurs dans $\mathbf{R}^d$: c'est un *champ de vecteurs*. Nous supposerons qu'en tout point $\mathbf{x}$ de $\partial\Omega$ est défini un *vecteur normal sortant* $\overrightarrow{n}(\mathbf{x})$ et que l'on peut choisir $\overrightarrow{n}(\mathbf{x})$ unitaire de sorte que $\mathbf{x} \mapsto \overrightarrow{n}(\mathbf{x})$ soit une application de classe $\mathcal{C}^1$. Le *flux du champ* $\overrightarrow{\mathrm{X}}$ *au travers de l'hypersurface* $\partial\Omega$ est alors donné par l'intégrale

$$\Phi_{\partial\Omega}(\overrightarrow{\mathrm{X}}) = \int_{\partial\Omega} \overrightarrow{n}(\mathbf{x}) \cdot \overrightarrow{\mathrm{X}}(\mathbf{x})\, \mathrm{d}\sigma(\mathbf{x}),$$

où $\mathrm{d}\sigma(\mathbf{x})$ dénote un « élément d'aire » sur l'hypersurface $\partial\Omega$. Cette intégrale est définie à partir des paramétrages locaux de $\partial\Omega$ introduits plus haut et, comme dans le cas de la formule de Green-Riemann, on montre que l'intégrale ne dépend pas du choix des paramétrages grâce à la formule du changement de variable.

Un exemple simple en dimension 3 : soit une surface paramétrée $\mathrm{S} \subset \mathbf{R}^3$ donnée comme l'image d'une application $\varphi : \mathrm{U} \longrightarrow \mathbf{R}^3$ qui soit injective, de classe $\mathcal{C}^1$ et de différentielle partout injective sur l'ouvert U de $\mathbf{R}^2$. Le paramétrage φ définit un vecteur normal à la nappe grâce au produit vectoriel des dérivée partielles

$$\overrightarrow{\mathrm{N}}_\varphi(u, v) = \frac{\partial\varphi}{\partial u}(u, v) \wedge \frac{\partial\varphi}{\partial v}(u, v).$$

L'élément d'aire $\mathrm{d}\sigma(\mathbf{x})$ sur la surface S est alors donné par $\mathrm{d}\sigma(\mathbf{x}) = \|\overrightarrow{\mathrm{N}}_\varphi(u, v)\|\, \mathrm{d}u\, \mathrm{d}v$. Le vecteur unitaire normal est donné par

$$\overrightarrow{n}(\mathbf{x}) = \overrightarrow{n}(\varphi(u, v)) = \left\|\overrightarrow{\mathrm{N}}_\varphi(u, v)\right\|^{-1} \overrightarrow{\mathrm{N}}_\varphi(u, v).$$

Si $\psi : \mathrm{V} \longrightarrow \mathbf{R}^3$ est un autre paramétrage de S, alors $\psi = \varphi \circ h^{-1}$ où $h : \mathrm{U} \longrightarrow \mathrm{V}$ est un difféomorphisme préservant l'orientation. Le paramétrage ψ définit lui aussi un vecteur normal à la nappe et les deux vecteurs normaux $\overrightarrow{\mathrm{N}}_\psi$ et $\overrightarrow{\mathrm{N}}_\varphi$ sont reliés par la formule

$$\overrightarrow{\mathrm{N}}_\varphi(u, v) = \mathrm{J}(h)(u, v) \overrightarrow{\mathrm{N}}_\psi(h(u, v)),$$

où $\mathrm{J}(h)$ est le jacobien de h. Si $\overrightarrow{\mathrm{X}}$ est un champ de vecteurs sur $\mathbf{R}^3$, son flux au travers de S est alors donné par

$$\Phi_{\mathrm{S}}(\overrightarrow{\mathrm{X}}) = \int_{\mathrm{S}} \overrightarrow{n}(\mathbf{x}) \cdot \overrightarrow{\mathrm{X}}(\mathbf{x})\, \mathrm{d}\sigma(\mathbf{x}) = \iint_{\mathrm{U}} \overrightarrow{n}(\varphi(u, v)) \cdot \overrightarrow{\mathrm{X}}(\varphi(u, v)) \left\|\overrightarrow{\mathrm{N}}_\varphi(u, v)\right\| \mathrm{d}u\, \mathrm{d}v$$

et la formule du changement de variable nous assure que nous trouverions le même résultat en prenant le paramétrage ψ.

Associée au champ de vecteur $\overrightarrow{\mathrm{X}}$ on trouve également sa *divergence*. C'est la fonction continue $\mathrm{div}(\overrightarrow{\mathrm{X}})$ sur $\overline{\Omega}$ définie par

$$\mathrm{div}(\overrightarrow{\mathrm{X}})(\mathbf{x}) = \sum_{k=1}^{d} \frac{\partial \mathrm{X}_k}{\partial x_k}(\mathbf{x}),$$

où l'on a écrit $\overrightarrow{\mathrm{X}}$ en coordonnées $\overrightarrow{\mathrm{X}} = (\mathrm{X}_1, \ldots, \mathrm{X}_k)$. Cette fonction est intégrable sur Ω et *la formule de la divergence* s'énonce alors :

$$\int_{\partial\Omega} \overrightarrow{n}(\mathbf{x}) \cdot \overrightarrow{\mathrm{X}}(\mathbf{x})\,\mathrm{d}\sigma(\mathbf{x}) = \int_{\Omega} \mathrm{div}(\overrightarrow{\mathrm{X}})(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Cette formule a de nombreuses applications en physique. Prenons un exemple en électrostatique. La première équation de Maxwell affirme que si $\overrightarrow{\mathrm{E}}$ est le champ électrique et ρ la densité de charges dans l'espace $\mathbf{R}^3$, alors

$$\mathrm{div}(\overrightarrow{\mathrm{E}}) = \frac{\rho}{\varepsilon_0},$$

où ε_0 est une constante fondamentale de la physique, la permittivité du vide. Soit maintenant Ω un ouvert borné de $\mathbf{R}^3$ délimité par un bord de classe $\mathcal{C}^\infty$. L'exemple le plus souvent utilisé est celui d'une boule, dont le bord n'est jamais qu'une sphère. Si l'on applique la formule de la divergence à cette situation, on trouve

$$\Phi_{\partial\Omega}(\overrightarrow{\mathrm{E}}) = \int_{\Omega} \mathrm{div}(\overrightarrow{\mathrm{E}})(\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{1}{\varepsilon_0}\int_{\Omega} \rho(\mathbf{x})\,\mathrm{d}\mathbf{x},$$

de sorte que le flux sortant de $\overrightarrow{\mathrm{E}}$ au travers de la frontière de Ω est proportionnel à la quantité de charge électrique totale contenue dans Ω. Cela permet, dans de nombreux cas, de calculer explicitement le champ $\overrightarrow{\mathrm{E}}$ en fonction de la densité de charges ρ.

10.7. Appendice : preuve de la formule de changement de variable

Nous allons ici donner la preuve de la formule de changement de variable pour le cas des fonctions différentiables à support compact. On obtient le cas général des fonctions Lebesgue-intégrables par l'application d'un théorème de densité dont nous verrons l'énoncé dans le prochain chapitre. En fait, toute la substance géométrique de la preuve se trouve déjà dans le cas régulier. Il existe de multiples preuves de la formule du jacobien. La plus naturelle consiste à découper le domaine d'intégration en petits cubes sur lesquels on peut approcher convenablement le difféomorphisme

de changement de variable par sa différentielle en un point, à appliquer dans ce cube la formule de changement de variable dans le cas linéaire puis à sommer sur tous les petits bouts, en espérant que les termes d'erreur restent petits après sommation. Une autre méthode consiste à effectuer une récurrence sur le nombre de variables affectées par le changement de variable en utilisant de manière astucieuse la formule de Fubini. Nous allons donner une preuve due au mathématicien Peter Lax en 1999 qui a l'avantage d'être relativement élémentaire mais le désavantage de ne devenir intuitive que si l'on connaît un petit peu la théorie du *calcul différentiel extérieur* qui traite des formes différentielles de degré plus grand que 1.

10.7.1. Le théorème de P. Lax

Soit f une fonction de classe $\mathcal{C}^1$ sur $\mathbf{R}^d$. Nous allons adopter une notation compacte et commode pour les dérivées partielles en posant

$$\frac{\partial f}{\partial x} = \partial_x f.$$

La différentielle de f au point $\mathbf{x}$ est la forme linéaire

$$\mathrm{D}f(\mathbf{x}) = \partial_{x_1} f(\mathbf{x})\,\mathrm{d}x_1 + \cdots + \partial_{x_d} f(\mathbf{x})\,\mathrm{d}x_d,$$

où nous avons repris la notation standard $\mathrm{d}x_i = \mathrm{e}_i^*$ pour le i^{e} élément de la base duale. Le gradient de f en $\mathbf{x}$ est le vecteur orthogonal à la forme $\mathrm{D}f(\mathbf{x})$ pour le produit scalaire euclidien usuel :

$$\vec{\nabla} f(\mathbf{x}) = (\partial_{x_1} f(\mathbf{x}), \ldots, \partial_{x_d} f(\mathbf{x}))$$

de sorte qu'appliquer $\mathrm{D}f(\mathbf{x})$ au vecteur $\vec{v}$ revient à prendre le produit scalaire de $\vec{v}$ avec $\vec{\nabla} f(\mathbf{x})$ [1].

Soit φ une application de classe $\mathcal{C}^1$ de $\mathbf{R}^d$ dans $\mathbf{R}^d$. Notons

$$\varphi(\mathbf{x}) = (\varphi_1(x_1, \ldots, x_d), \ldots, \varphi_d(x_1, \ldots, x_d))$$

de sorte que la *matrice jacobienne* de φ en $\mathbf{x}$ soit la matrice carrée

$$\mathrm{M}(\mathbf{x}) = \left(\frac{\partial \varphi_j}{\partial x_i}(\mathbf{x})\right)_{1 \leqslant i,j \leqslant d} = (\partial_{x_i} \varphi_j(\mathbf{x}))_{1 \leqslant i,j \leqslant d}.$$

Les vecteurs colonnes de la jacobienne ne sont autres que les gradients des composantes de φ. Nous noterons $\mathrm{J}(\mathbf{x})$ le *jacobien* de φ en $\mathbf{x}$, c'est-à-dire le déterminant de la matrice $\mathrm{M}(\mathbf{x})$. Nous pouvons maintenant énoncer le théorème de Lax.

Théorème 10.11 — *Soient f une fonction de classe $\mathcal{C}^1$ définie sur $\mathbf{R}^d$ et nulle hors d'une boule de rayon assez grand. Soit φ une application de classe $\mathcal{C}^2$ de $\mathbf{R}^d$ dans*

1. Le symbole $\vec{\nabla}$ s'appelle *nabla*.

lui-même telle que, pour tout $\mathbf{x}$ *hors de la boule unité, on ait* $\varphi(\mathbf{x}) = \mathbf{x}$*. Alors la fonction* $\mathbf{x} \mapsto f(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x})$ *est intégrable sur* $\mathbf{R}^d$ *et*

$$\int_{\mathbf{R}^d} f(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbf{R}^d} f(\mathbf{y})\,\mathrm{d}\mathbf{y}.$$

Démonstration — Notons $r > 0$ un rayon pour lequel $f(\mathbf{x}) = 0$ si $\|\mathbf{x}\| > r$. Nous avons tout loisir de supposer r plus grand que 1, ce que nous ferons. Comme $\varphi(\mathbf{x}) = \mathbf{x}$ pour tout $\mathbf{x}$ de norme plus grande que 1, on en déduit que si $\|\mathbf{x}\| > r$ on a $f(\varphi(\mathbf{x})) = f(\mathbf{x}) = 0$. La fonction $\mathbf{x} \mapsto f(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x})$ est donc nulle hors de la boule de rayon r. Comme elle est continue, elle est intégrable et la première partie du théorème est démontrée. Reste à prouver la formule !

Nous noterons dans la suite $\mathrm{C}_r = [-r, r] \times \cdots \times [-r, r]$ le cube d-dimensionnel de rayon r centré en 0. Définissons la fonction g sur $\mathbf{R}^d$ par

$$g(y_1, y_2, \ldots, y_d) = \int_{-\infty}^{y_1} f(z, y_2, \ldots, y_d)\,\mathrm{d}z.$$

Elle est bien définie du fait que f est nulle hors de C_r et l'on peut de plus affirmer que si $|y_k| > r$ pour un certain $k \geqslant 2$ ou si $y_1 \leqslant -r$, on a $g(y_1, \ldots, y_d) = 0$. La fonction g ne prend donc des valeurs non nulles que dans la bande

$$\mathrm{B}_r = [-r, +\infty[\times [-r, r] \times \cdots \times [-r, r].$$

Par définition de g on a bien évidemment

$$\partial_{y_1} g = f,$$

de sorte que le premier membre de l'égalité énoncée dans le théorème peut s'écrire

$$\int_{\mathbf{R}^d} f(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbf{R}^d} \frac{\partial g}{\partial y_1}(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Mais f étant nulle hors du cube, nous pouvons en fait limiter les domaines d'intégration dans ces deux intégrales au cube C_r.

D'après la règle de dérivation des fonctions composées, nous avons

$$\overrightarrow{\nabla}\, g \circ \varphi(\mathbf{x}) = \sum_{j=1}^{d} \partial_{y_j} g(\varphi(\mathbf{x}))\,\overrightarrow{\nabla}\,\varphi_j(\mathbf{x})$$

de sorte que $\overrightarrow{\nabla}\, g \circ \varphi$ est une combinaison linéaire des $\overrightarrow{\nabla}\,\varphi_j$. Comme un déterminant ne change pas par opérations élémentaires sur les colonnes, nous obtenons que

$$\begin{aligned}\det\big(\overrightarrow{\nabla}\, g \circ \varphi(\mathbf{x}), \overrightarrow{\nabla}\,\varphi_2(\mathbf{x}), \ldots, \overrightarrow{\nabla}\,\varphi_d(\mathbf{x})\big)\\ = \det\big(\partial_{y_1} g(\varphi(\mathbf{x}))\,\overrightarrow{\nabla}\,\varphi_1(\mathbf{x}), \overrightarrow{\nabla}\,\varphi_2(\mathbf{x}), \ldots, \overrightarrow{\nabla}\,\varphi_d(\mathbf{x})\big).\end{aligned}$$

En sortant le terme $\partial_{y_1} g(\varphi(\mathbf{x}))$ par multilinéarité du déterminant, nous en tirons l'identité suivante :

$$f(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \frac{\partial g}{\partial y_1}(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x}) = \det\big(\overrightarrow{\nabla}\, g\circ\varphi(\mathbf{x}), \overrightarrow{\nabla}\,\varphi_2(\mathbf{x}), \ldots, \overrightarrow{\nabla}\,\varphi_d(\mathbf{x})\big).$$

Nous pouvons développer le déterminant dans le troisième membre ci-dessus selon la première colonne ce qui, en notant $\mathrm{M}_1(\mathbf{x}), \ldots, \mathrm{M}_d(\mathbf{x})$ les cofacteurs de la première colonne de la jacobienne, donne

$$\det\big(\overrightarrow{\nabla}\, g\circ\varphi(\mathbf{x}), \overrightarrow{\nabla}\,\varphi_2(\mathbf{x}), \ldots, \overrightarrow{\nabla}\,\varphi_d(\mathbf{x})\big) = \sum_{j=1}^{d} \mathrm{M}_j(\mathbf{x})\partial_{x_j} g(\varphi(\mathbf{x})).$$

Nous en déduisons que

$$\int_{\mathbf{R}^d} f(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathrm{C}_r} f(\varphi(\mathbf{x}))\mathrm{J}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathrm{C}_r}\left(\sum_{j=1}^{d} \mathrm{M}_j(\mathbf{x})\partial_{x_j} g(\varphi(\mathbf{x}))\right)\mathrm{d}\mathbf{x}.$$

Si $k \geqslant 2$, l'intégrale du k^{e} terme dans la somme ci-dessus s'écrit, d'après la formule de Fubini :

$$\int_{\mathrm{C}'_r}\left(\int_{-r}^{r} \mathrm{M}_k(\mathbf{x})\partial_{x_k} g(\varphi(\mathbf{x}))\,\mathrm{d}x_k\right)\mathrm{d}\mathbf{x}',$$

où C'_r est le carré unité de rayon r dans l'espace $\mathbf{R}^{d-1}$ des variables $\mathbf{x}' = (x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_d)$. Or une intégration par parties permet d'estimer l'intégrale d'une variable x_k, qui devient alors

$$\begin{aligned}\int_{-r}^{r} \mathrm{M}_k(\mathbf{x})\partial_{x_j} g(\varphi(\mathbf{x}))\,\mathrm{d}x_k = {} & [\mathrm{M}_k \cdot g\circ\varphi](x_1, \ldots, x_{k-1}, r, x_{k+1}, \ldots, x_d) \\ & - [\mathrm{M}_k \cdot g\circ\varphi](x_1, \ldots, x_{k-1}, -r, x_{k+1}, \ldots, x_d) \\ & - \int_{-r}^{r} \partial_{x_k}\mathrm{M}_k(\mathbf{x}) g(\varphi(\mathbf{x}))\,\mathrm{d}x_k.\end{aligned}$$

Comme $g(\mathbf{x}) = \mathbf{x}$ dès qu'une des coordonnées $x_2, \ldots, x_d$ dépasse r en valeur absolue, le premier terme du second membre est nul et il n'en reste que le terme intégral.

Le même raisonnement appliqué au premier terme de la somme écrite plus haut permet d'écrire celui-ci sous la forme

$$\int_{\mathrm{C}'_r}\left(\int_{-r}^{r} \mathrm{M}_1(\mathbf{x})\partial_{x_1} g(\varphi(\mathbf{x}))\,\mathrm{d}x_1\right)\mathrm{d}\mathbf{x}',$$

où C'_r et $\mathbf{x}'$ ont un sens évident. Une intégration par parties donne alors

$$\begin{aligned}\int_{-r}^{r} \mathrm{M}_1(\mathbf{x})\partial_{x_1} g(\varphi(\mathbf{x}))\,\mathrm{d}x_1 = {} & \mathrm{M}_1(r, x_2, \ldots, x_d) g(\varphi(r, x_2, \ldots, x_d)) \\ & - \mathrm{M}_1(-r, x_2, \ldots, x_d) g(\varphi(-r, x_2, \ldots, x_d)) \\ & - \int_{-r}^{r} \partial_{x_1}\mathrm{M}_1(\mathbf{x}) g(\varphi(\mathbf{x}))\,\mathrm{d}x_1.\end{aligned}$$

Mais φ est égal à l'identité hors du carré et donc $\mathrm{M}_1(r, x_2, \ldots, x_d) = \mathrm{M}_1(-r, x_2, \ldots, x_d) = 1$. Cela entraîne aussi que $g(\varphi(-r, x_2, \ldots, x_d)) = 0$ et $g(\varphi(r, x_2, \ldots, x_d)) = g(r, x_2, \ldots, x_d)$, de sorte que finalement

$$\int_{-r}^{r} \mathrm{M}_1(\mathbf{x}) \partial_{x_1} g(\varphi(\mathbf{x}))\, \mathrm{d}x_1 = g(r, x_2, \ldots, x_d) - \int_{-r}^{r} \partial_{x_1} \mathrm{M}_1(\mathbf{x}) g(\varphi(\mathbf{x}))\, \mathrm{d}x_1.$$

Rassemblons tout cela dans les intégrales multiples que nous avons cassées avec la formule de Fubini et nous trouvons

$$\int_{\mathbf{R}^d} f(\varphi(\mathbf{x})) \mathrm{J}(\mathbf{x})\, \mathrm{d}\mathbf{x} = \int_{\mathrm{C}'_r} g(r, x_2, \ldots, x_d)\, \mathrm{d}\mathbf{x}' - \int_{\mathrm{C}_r} g(\varphi(\mathbf{x})) \sum_{k=1}^{d} \partial_{x_k} \mathrm{M}_k(\mathbf{x})\, \mathrm{d}\mathbf{x}.$$

Cependant, par définition de g et grâce encore une fois à la formule de Fubini, nous avons

$$\int_{\mathrm{C}'_r} g(r, x_2, \ldots, x_d)\, \mathrm{d}\mathbf{x}' = \int_{\mathrm{C}'_r} \int_{-r}^{r} f(x_1, x_2, \ldots, x_d)\, \mathrm{d}x_1\, \mathrm{d}\mathbf{x}' = \int_{\mathbf{R}^d} f(\mathbf{x})\, \mathrm{d}\mathbf{x}.$$

Il nous suffit donc, pour parachever la preuve du théorème de Lax, de démontrer le lemme suivant :

Lemme 10.12 — *Avec les notations introduites ci-avant, on a, pour tout* $\mathbf{x} \in \mathbf{R}^d$,

$$\sum_{k=1}^{d} \partial_{x_k} \mathrm{M}_k(\mathbf{x}) = 0.$$

Nous laissons au lecteur le soin de démontrer ce lemme, en commençant par le cas de 2 variables, où il est une conséquence simple de l'identité de Schwarz $\partial_{x_1}\partial_{x_2}\varphi_2 = \partial_{x_2}\partial_{x_1}\varphi_2$. Pour le lecteur qui connaît un peu de calcul différentiel extérieur, le lemme est conséquence simple du fait que

$$\mathrm{d}(\,\mathrm{d}\varphi_2 \wedge \cdots \wedge \mathrm{d}\varphi_d) = \left(\sum_{k=1}^{d} \partial_{x_k} \mathrm{M}_k \right) \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_d$$

et de la relation $d \circ d = 0$. □

Avant de passer à la suite de notre programme, notons que dans la formule de Lax le fait que φ soit inversible n'intervient pas et que l'on ne prend pas la valeur absolue du jacobien. Le deuxième fait s'explique assez facilement : si φ est un difféomorphisme, son jacobien ne peut s'annuler et définit une fonction continue sur $\mathbf{R}^d$. Comme il vaut $+1$ hors d'une boule, il est partout strictement positif : les difféomorphismes qui satisfont aux hypothèses du théorème de Lax préservent l'orientation.

10.7.2. Application : le théorème du point fixe de Brouwer

La formule de changement de variable de Lax a de nombreuses applications. Nous allons ici montrer comment nous pouvons en déduire un important théorème de topologie, le théorème du point fixe de Brouwer, dont voici l'énoncé :

Théorème 10.13 — *Soit ψ une application continue de la boule unité fermée de $\mathbf{R}^d$ dans elle-même. Alors ψ admet un point fixe.*

Démonstration — Commençons par démontrer le lemme suivant :

Lemme 10.14 — *Si φ est une application qui satisfait aux hypothèses du théorème de Lax, alors φ est surjective.*

Démonstration — Supposons que φ n'est pas surjective et prenons donc un point $\mathbf{y}_0$ de $\mathbf{R}^d$ qui ne soit pas dans l'image de φ. Comme φ est égale à l'identité hors de la boule unité, le point $\mathbf{y}_0$ est nécessairement dans icelle. Mais par ailleurs l'image de la boule unité est un fermé puisque φ est continue et donc il existe une petite boule ouverte non vide B_0 centrée en $\mathbf{y}_0$ dans laquelle il n'y a aucune image de φ. Prenons maintenant une fonction f qui soit nulle hors de B_0 et d'intégrale non nulle. La formule de Lax nous permet alors d'affirmer que

$$\int_{\mathbf{R}^d} f(\mathbf{y})\,\mathrm{d}\mathbf{y} = \int_{\mathbf{R}^d} f(\varphi(\mathbf{x}))\,\mathrm{J}(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Le premier terme est nul du fait que B_0 ne contient aucune image de φ tandis que le second terme est non nul par hypothèse. C'est une absurdité et le lemme est démontré.

□

Le théorème du point fixe de Brouwer traite des applications continues de la boule unité dans elle-même tandis que le théorème de Lax nous parle de changements de variable qui sont de classe $\mathcal{C}^2$. Le lemme suivant nous dit que les premières peuvent s'approcher par les seconds :

Lemme 10.15 — *Soient $\mathbf{B}_d$ la boule unité fermée de $\mathbf{R}^d$ et $\varphi : \mathbf{B}_d \longrightarrow \mathbf{R}^d$ une application continue qui soit égale à l'identité à la frontière : $\forall \mathbf{x} \in \partial\mathbf{B}_d, \varphi(\mathbf{x}) = \mathbf{x}$. Alors il existe une suite d'applications (φ_n) de classe $\mathcal{C}^2$ de $\mathbf{R}^d$ dans lui-même, telles que φ_n soit égal à l'identité hors de $\mathbf{B}_d$ et que φ soit limite uniforme de la suite (φ_n) dans $\mathbf{B}_d$.*

Nous admettrons ce lemme, qui est une conséquence simple du théorème de Stone-Weierstraß : toute fonction continue sur un segment $[a, b]$ est limite uniforme d'une suite de fonctions polynomiales. Nous pouvons maintenant démontrer le lemme suivant qui est une sorte d'analogue multidimensionnel du théorème des valeurs intermédiaires :

Lemme 10.16 — *Soit $\varphi : \mathbf{B}_d \longrightarrow \mathbf{R}^d$ une application continue qui soit égale à l'identité à la frontière. Alors $\mathbf{B}_d \subset \varphi(\mathbf{B}_d)$.*

Considérons en effet la suite (φ_n) d'applications de classe $\mathcal{C}^2$ donnée par le lemme 10.15. Pour tout $n \geqslant 0$, φ_n est surjective d'après le lemme 10.14. Fixons

maintenant $\mathbf{y} \in \mathbf{B}_d$. Pour chaque $n \in \mathbf{N}$ il existe $\mathbf{x}_n \in \mathbf{B}_d$ tel que $\varphi_n(\mathbf{y}_n) = \mathbf{x}_n$. Comme $\mathbf{B}_d$ est compacte, il existe une suite $(\mathbf{x}_{n_k})_{k\in\mathbf{N}}$ extraite de la suite $(\mathbf{x}_n)$ qui converge vers un point $\mathbf{x}$ de $\mathbf{B}_d$. Le point $\mathbf{x}$ satisfait clairement à $\varphi(\mathbf{x}) = \mathbf{y}$ et cela achève la démonstration du lemme 10.16.

Nous pouvons terminer la preuve du théorème de Brouwer. Soit $\psi : \mathbf{B}_d \longrightarrow \mathbf{B}_d$ une application continue. Raisonnons par l'absurde et supposons que ψ ne possède pas de point fixe. Pour tout $\mathbf{x} \in \mathbf{B}_d$ on peut alors mener la demi-droite partant de $\psi(\mathbf{x})$ et passant par $\mathbf{x}$. Cette demi-droite coupe la sphère unité $\partial\mathbf{B}_d$ en un point que nous noterons $\varphi(\mathbf{x})$. L'application φ est continue et satisfait, par construction, à $\varphi(\mathbf{B}_d) \subset \partial\mathbf{B}_d$ mais aussi à

$$\forall \mathbf{x} \in \mathbf{B}_d, \varphi(\mathbf{x}) = \mathbf{x}.$$

Cela contredit le lemme 10.16 et nous avons terminé la preuve du théorème de Brouwer. □

Remarquons pour finir que le lemme 10.15 permet d'étendre la formule de Lax au cas où le changement de variable φ est juste supposé de classe $\mathcal{C}^1$.

10.7.3. Changement de variable : la formule classique

Rappelons pour commencer la formule du changement de variable telles que nous l'avons énoncée dans ce chapitre.

Théorème 10.17 — *Soient Ω un ouvert de $\mathbf{R}^d$, $\varphi : \Omega \longrightarrow \varphi(\Omega) \subset \mathbf{R}^d$ un difféomorphisme de classe $\mathcal{C}^1$ et $f : \varphi(\Omega) \longrightarrow \mathbf{R}$ une fonction Lebesgue-intégrable. Alors $f \circ \varphi$ est Lebesgue-intégrable sur Ω et on a la formule*

$$\int_{\varphi(\Omega)} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\Omega} f \circ \varphi(\mathbf{u}) |(\mathrm{Jac}\,\varphi)(\mathbf{u})|\,\mathrm{d}\mathbf{u}.$$

Cinq différences avec le théorème de Lax sautent aux yeux : tout d'abord f est supposée Lebesgue-intégrable et non $\mathcal{C}^1$, φ n'est pas définie sur tout $\mathbf{R}^d$ mais est un difféomorphisme de Ω sur son image $\varphi(\Omega)$, φ est de classe $\mathcal{C}^1$ et non $\mathcal{C}^2$ et enfin le jacobien de φ est habillé de valeurs absolues.

Le passage des fonctions de classe $\mathcal{C}^1$ aux fonctions Lebesgue-intégrable est une conséquence simple des théorèmes de densité que nous démontrerons dans le chapitre suivant. Ensuite le fait que φ soit de classe $\mathcal{C}^1$ et non $\mathcal{C}^2$ peut être aisément contourné grâce au théorème de Stone-Weierstraß, comme nous l'avons fait pour le lemme 10.15 : un difféomorphisme de classe $\mathcal{C}^1$ est limite uniforme de difféomorphismes de classes $\mathcal{C}^2$ sur des compacts. Mais ce qui va réellement poser problème, c'est de pouvoir étendre φ en une application qui vaille l'identité en dehors d'une grande boule, tout en respectant suffisamment les domaines de fonctions que l'on intègre. Ici la géométrie intervient et les difficultés surgissent !

La formule de Lax est souvent présentée comme une généralisation de la formule classique, cette dernière en étant présentée comme une conséquence simple. Il n'en

est cependant rien et les approches naïves mènent immanquablement à des erreurs. Pour démontrer proprement la formule du jacobien, nous avons besoin de procédures d'approximation et de localisation qui sont très classiques en analyse. Nous les présentons ci-dessous sans entrer dans les détails.

Partitions de l'unité

Rappelons que le *support* d'une fonction continue est, par définition, l'adhérence de l'ensemble des points où cette fonction est non nulle. C'est un fermé hors duquel la fonction, comme il se doit, est nulle. Si f est une fonction définie sur un ouvert Ω' (égal, dans notre cas, à $\varphi(\Omega)$), il peut être intéressant de casser Ω' en petits bouts et de décomposer f en une somme de fonctions définies sur ces petits bouts. Cela est bien évidemment toujours possible si l'on n'impose rien aux restrictions de f aux petits morceaux de Ω'. Si l'on veut leur imposer, par exemple, la continuité, il faut être soigneux. Nous allons faire cela dans le cas où l'on décompose Ω' en une union de boules ouvertes. Comme $\mathbf{Q}^d$ est dense dans $\mathbf{R}^d$, il n'est pas difficile de démontrer que si $r > 0$, il existe une famille *dénombrable* de boules ouvertes $\mathrm{B}_i = \mathrm{B}_i(\mathbf{x}_i, r_i), i \in \mathbf{N}$ telle que pour tout i on ait $\mathrm{B}_i \subset \Omega'$, $0 < r_i \leqslant r$ et telle que les B_i recouvrent Ω' :

$$\Omega' = \bigcup_{i\in\mathbf{N}} \mathrm{B}_i.$$

On peut de plus imposer à ce recouvrement par des boules d'être *localement fini :* tout point $\mathbf{x}$ de Ω' est dans un nombre au plus fini de boules du recouvrement. Nous aimerions maintenant construire une famille de fonctions continues $f_i, i \in \mathbf{N}$ telles que

– pour tout i, f_i est de support inclus dans B_i ;

– les f_i forment une *partition de la fonction* f au sens où

$$\forall \mathbf{x} \in \Omega', \sum_{i=1}^{\infty} f_i(\mathbf{x}) = f(\mathbf{x}).$$

Cela peut aisément être construit. Si $\mathbf{x} \in \Omega'$ et $i \in \mathbf{N}$, la distance de $\mathbf{x}$ au complémentaire de la boule B_i vaut par définition 0 si $\mathbf{x} \notin \mathrm{B}_i$ et $r - \|\mathbf{x} - \mathbf{x}_i\|$ si $\mathbf{x} \in \mathrm{B}_i$. Notons d_i la fonction ainsi définie : d_i est continue sur Ω' et son support est inclus dans B_i. Par ailleurs pour tout $\mathbf{x} \in \mathbf{N}$, les indices i pour lesquels $d_i(\mathbf{x})$ est non nul sont en nombre fini et non nul, de sorte que la fonction

$$\mathbf{x} \longmapsto \sum_{i=1}^{\mathbf{N}} d_i(\mathbf{x})$$

est bien définie, continue, et partout non nulle. On peut alors poser, pour tout $\mathbf{x} \in \Omega'$ et tout $i \in \mathbf{N}$,

$$p_i(\mathbf{x}) = \frac{d_i(\mathbf{x})}{\displaystyle\sum_{i=1}^{\infty} d_i(\mathbf{x})}$$

La famille des fonctions p_i satisfait aux deux propriétés suivantes :

– pour tout i, p_i est continue et de support inclus dans B_i ;
– les p_i forment une *partition de l'unité* subordonnée au recouvrement de Ω' par la famille de boules (B_i) :

$$\forall \mathbf{x} \in \Omega', \quad \sum_{i=0}^{\infty} p_i(\mathbf{x}) = 1.$$

Nous pouvons donc maintenant poser $f_i(\mathbf{x}) = p_i(\mathbf{x}) f(\mathbf{x})$ et obtenir ainsi une partition de la fonction f comme indiqué plus haut.

On peut raffiner la construction ci-dessus et montrer que l'on peut en plus imposer aux fonctions p_i d'être de classe $\mathcal{C}^\infty$ et de support compact dans la boule ouverte B_i. Nous n'aurons pas besoin d'un tel résultat ici. Remarquons également que la construction donnée fonctionne si l'on remplace les boules B_i par des ouverts quelconques.

Muni de l'outil des partitions de l'unité, la linéarité de l'intégrale et le théorème de convergence dominée permettent de réduire l'énoncé au cas où le support de f est inclus dans une boule B de rayon petit et où φ est un difféomorphisme de $\varphi^{-1}(\mathrm{B})$ sur B. Le problème maintenant est d'étudier la géométrie de ce difféomorphisme sur une telle boule afin d'être capables de le prolonger à $\mathbf{R}^d$ pour pouvoir appliquer la formule de Lax. Comme nous l'avons noté, il est nécessaire dans ce cas que φ préserve l'orientation. Nous allons faire cette hypothèse et verrons plus tard comment s'en débarrasser.

Déformation d'un difféomorphisme local

Soient B une boule de rayon $\varepsilon > 0$ et φ un difféomorphisme de classe $\mathcal{C}^2$ défini sur un ouvert Ω, de jacobien positif et tel que $\varphi(\Omega)$ contienne $\overline{\mathrm{B}}$. Soit par ailleurs f une fonction continue à support compact inclus dans B. Pour appliquer la formule de Lax, nous devons construire une application ψ de classe $\mathcal{C}^1$ sur $\mathbf{R}^d$ telle que

– $\psi = \varphi$ sur $\varphi^{-1}(\mathrm{B})$;
– $\psi^{-1}(\mathrm{B}) = \varphi^{-1}(\mathrm{B})$;
– et ψ est égale à l'identité hors d'une boule assez grande,

le tout quitte à diminuer ε si nécessaire. L'idée pour parvenir à nos fins est simple : on applique à φ la formule de Taylor de sorte que φ est approximativement égal, au voisinage du point $\mathbf{x}_0$, à l'application affine $\varphi(\mathbf{x}_0) + \mathrm{D}\varphi(\mathbf{x}_0)$. On effectue ensuite une interpolation entre l'application linéaire $\mathrm{D}\varphi(\mathbf{x}_0)$ et l'identité à l'aide du lemme suivant :

Lemme 10.18 — *Soit* M *une matrice inversible de déterminant positif d. Il existe une application h de classe $\mathcal{C}^\infty$ de* $[0,1]$ *dans* $\mathrm{GL}(d, \mathbf{R})$ *telle que* $h(0) = \mathrm{M}$, $h(1) = \mathrm{Id}$ *et* $\det(h(t))$ *est compris entre* 1 *et d pour tout t.*

Nous laissons au lecteur masochiste le soin de rédiger les détails de la preuve, consistant à soigneusement recoller φ sur une petite boule à l'identité hors d'une autre petite

boule de même centre en utilisant la déformation h entre $\mathrm{D}\varphi(\mathbf{x}_0)$ et Id donnée par le lemme. Donnons-en une indication. Pour simplifier, supposons $\mathbf{x}_0 = 0$ et $\varphi(\mathbf{x}_0) = 0$. On peut trouver un rayon $0 < \eta < \varepsilon$ tel que l'application ψ définie comme suit convienne :

- $\psi(\mathbf{x}) = \varphi(\mathbf{x})$ si $\|\mathbf{x}\| \leqslant \eta$;
- $\psi(\mathbf{x}) = \mathrm{D}\varphi(\mathbf{x}_0)(\mathbf{x}) + \alpha(\|x\|)[\varphi(\mathbf{x}) - \mathrm{D}\varphi(\mathbf{x}_0)(\mathbf{x})]$ pour $\eta \leqslant \|\mathbf{x}\| \leqslant 2\eta$, où α est une fonction décroissante de classe $\mathcal{C}^2$ valant 1 en η et 0 en 2η ;
- $\psi(\mathbf{x}) = h(\|\mathbf{x}\|/\eta - 2)$ si $2\eta \leqslant \|\mathbf{x}\| \leqslant 3\eta$, où h est la déformation donnée par le lemme ci-dessus ;
- $\psi(\mathbf{x}) = \mathbf{x}$ si $\|\mathbf{x}\| \geqslant 3\eta$.

Il nous reste finalement à traiter le cas où φ ne préserve pas l'orientation. On se ramène au cas préservant l'orientation en précomposant φ par une application linéaire simple et renversant l'orientation, par exemple $\sigma(x_1, \ldots, x_d) = (x_2, x_1, x_3, \ldots, x_d)$, dont le déterminant vaut -1. Le difféomorphisme $f \circ \sigma$ préserve l'orientation et il suffit alors de démontrer la formule de changement de variable dans le cas du changement de variable linéaire σ. Cela est une conséquence facile de la formule de Fubini. En effet celle-ci nous dit que

$$\begin{aligned}
\int_{\mathbf{R}^d} f(\mathbf{x})\,\mathrm{d}\mathbf{x} &= \int_{\mathbf{R}}\int_{\mathbf{R}}\int_{\mathbf{R}^{d-2}} f(x_1, x_2, x_3, \ldots, x_d)\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}\mathbf{x}' \\
&= \int_{\mathbf{R}^{d-2}} \left(\int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x_1, x_2, x_3, \ldots, x_d)\,\mathrm{d}x_1 \right) \mathrm{d}x_2 \right) \mathrm{d}\mathbf{x}' \\
&= \int_{\mathbf{R}^{d-2}} \left(\int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x_2, x_1, x_3, \ldots, x_d)\,\mathrm{d}x_1 \right) \mathrm{d}x_2 \right) \mathrm{d}\mathbf{x}' \\
&= \int_{\mathbf{R}^d} f(\sigma(\mathbf{x}))|\det(\sigma)|\,\mathrm{d}\mathbf{x}.
\end{aligned}$$

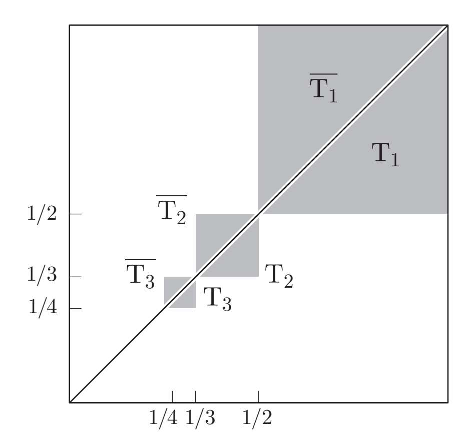
10.7.4. Un cas pathologique pour le changement de variable

Comme nous l'avons déjà mentionné, les intégrales semi-convergentes se comportent mal quand on intègre sur un domaine de dimension d plus grande que 2, ce qui rend leur définition même problématique : nous allons donner un exemple d'une fonction f intégrable (mais pas Lebesgue-intégrable) sur $\mathbf{R}^2$ telle qu'un changement de variable linéaire rende la fonction non intégrable. Comme une théorie cohérente de l'intégration se doit d'être indépendante du choix d'un système de coordonnées, un tel phénomène montre que si l'on autorise des intégrales semi-convergentes, on perd cette invariance, et cela justifie notre choix fait dans ce chapitre de ne parler que de fonctions Lebesgue-intégrables.

Pour un entier $n \geqslant 1$, on note T_n le triangle de $\mathbf{R}^2$ défini par

$$\mathrm{T}_n = \left\{ (x, y) \in \mathbf{R}^2, \quad \frac{1}{n+1} < x < \frac{1}{n} \ \text{ et } \ \frac{1}{n+1} < y < x \right\}.$$

Les triangles ouverts T_n sont tous situés au-dessous de la diagonale du carré unité $\mathrm{P} = [0,1] \times [0,1]$. Soit $\overline{\mathrm{T}_n}$ le symétrique de T_n par rapport à la diagonale. Nous obtenons alors la figure suivante :



On définit alors la fonction $f : \mathbf{R}^2 \longrightarrow \mathbf{R}$ de la manière suivante :

$$\begin{cases} f(x,y) = n(n+1)^2 \ \text{ si } \ (x,y) \in \mathrm{T}_n \\ f(x,y) = -n(n+1)^2 \ \text{ si } \ (x,y) \in \overline{\mathrm{T}_n} \\ f(x,y) = 0 \ \text{ sinon.} \end{cases}$$

La fonction f est intégrable (au sens de Kurzweil-Henstock) mais pas Lebesgue intégrable. Cela est dû, d'une part à la convergence de la série harmonique alternée $\sum(-1)^n n^{-1}$ et d'autre-part au fait que nous avons construit notre fonction par morceaux *en nous appuyant sur les axes de coordonnées.* Si l'on « tourne le dessin » ci-dessus de quarante-cinq degrés, la fonction obtenue n'est plus intégrable. Soit φ l'isomorphisme linéaire de $\mathbf{R}^2$ dans lui-même dont la matrice dans la base canonique est donnée par

$$\mathrm{A} = \begin{pmatrix} 2^{-1/2} & -(2^{-1/2}) \\ 2^{-1/2} & 2^{-1/2} \end{pmatrix},$$

ce qui fait que φ n'est autre que la rotation d'angle $\pi/4$, dont le jacobien vaut 1. Sur le carré $[0, \sqrt{2}] \times [-\sqrt{2}, 0]$, la fonction $f \circ \varphi$ n'est clairement pas intégrable, car dans ce carré on n'a pas de compensation des signes dans la série harmonique et l'intégrale que l'on obtient vaut $\sum n^{-1} = +\infty$. Or il n'est pas difficile de démontrer que si une fonction (ici $f \circ \varphi$) est intégrable sur un pavé P (ici par exemple $[-2,2] \times [-2,2]$) qui contient $\varphi^{-1}([0,1] \times [0,1])$ alors elle est intégrable sur tout pavé inclus dans P. La fonction $f \circ \varphi$ ne peut donc être intégrable, et l'on ne peut donc avoir de formule de changement de variable générale pour les fonctions intégrables au sens de Kurzweil-Henstock, du moins de théorème où l'existence de l'intégrale de $f \circ \varphi$ entraîne l'existence de celle de f. Ces questions sont un sujet de recherche actif touchant aux liens entre théorie de l'intégration et géométrie : la *théorie géométrique de la mesure.*

Chapitre 11

Mesure de Lebesgue, espaces L^p, applications

L'une des idées les plus fécondes de l'analyse moderne a été de considérer les fonctions comme des points dans certains espaces sur lesquels utiliser l'intuition géométrique. Ainsi, si $\mathrm{K} \subset \mathbf{R}^d$ est un compact, l'espace $\mathcal{C}^0(\mathrm{K}, \mathbf{R})$ des fonctions numériques continues sur K est naturellement muni d'une structure de $\mathbf{R}$-espace vectoriel (et même de $\mathbf{R}$-algèbre). Cet espace n'est de dimension finie que si K est un ensemble fini. Dans tous les cas on peut munir $\mathcal{C}^0(\mathrm{K}, \mathbf{R})$ de la norme uniforme

$$\|f\|_\infty = \sup\left\{|f(\mathbf{x})|,\ \mathbf{x} \in \mathrm{K}\right\}$$

et cela fait de $\mathcal{C}^0(\mathrm{K}, \mathbf{R})$ un espace vectoriel normé et complet : on dit que c'est un espace de Banach. On peut en étudier les propriétés algébriques et géométriques comme pour n'importe quel espace vectoriel normé, et il se trouve qu'il existe un dictionnaire entre les propriétés de cet espace et les propriétés topologiques de K. Le théorème de Gelfand, donné en exercice et décrivant les idéaux de $\mathcal{C}^0(\mathrm{K}, \mathbf{R})$ en est un exemple fondamental, qui illustre l'intérêt d'avoir de jolis espaces fonctionnels sur lesquels travailler.

Dans le cadre de la théorie de l'intégration il ressort une famille d'espaces très importants, les espaces de Lebesgue dits « espaces L^p » ; il y en a un pour chaque réel $p \geqslant 1$, et l'espace L^1 correspond aux fonctions Lebesgue-intégrables. Le but de ce chapitre est de donner une définition propre de ces espaces et de démontrer leur caractère complet pour la norme qui leur est naturellement associée. C'est, après le théorème de convergence dominée, un des résultats fondamentaux de l'analyse moderne.

11.1. Fonctions mesurables

Nous présentons dans ce livre la théorie de l'intégrale de Lebesgue qui est une théorie de l'intégration sur $\mathbf{R}^d$ qui possède la propriété d'être invariante par translation et

d'intégrer les fonctions en escalier. Nous nous sommes pour cela basés sur la notion de *volume* d-dimensionnel d'un pavé. Le théorème de convergence dominée va nous permettre d'étendre la notion de volume, qui prendra alors le nom de *mesure* de Lebesgue, à une vaste classe de parties de $\mathbf{R}^d$: les parties *mesurables* (au sens de Lebesgue).

Nous aurions tout aussi bien pu procéder de manière inverse et définir d'abord la classe des parties mesurables, puis leur mesure et enfin, par des passages à la limite utilisant des combinaisons linéaires de fonctions indicatrices de parties mesurables, définir l'intégrabilité et l'intégrale de fonctions variées. Cette approche a l'avantage d'être généralisable à bien d'autres contextes que l'intégration sur l'espace $\mathbf{R}^d$. On y considère un ensemble X, muni d'une famille de parties $\mathcal{M}$ que l'on décrète mesurables et d'une fonction réelle μ définie sur cette famille de parties de sorte que $\mu(\mathrm{A})$ représente la mesure de la partie mesurable $\mathrm{A} \in \mathcal{M}$. Sous de bonnes hypothèses sur $\mathcal{M}$ et μ (nous les énonçons plus loin dans le cadre de la mesure de Lebesgue sur $\mathbf{R}^d$ qui nous intéresse ici : $\mathcal{M}$ doit en particulier être ce que l'on appelle une *tribu*) on peut construire à partir de cela une théorie de l'intégration qui satisfait aux théorèmes de convergence de l'intégrale de Lebesgue. Une fois la théorie générale mise en place, il faut l'incarner et construire, par exemple, la mesure de Lebesgue. Hormis le volume, défini pour les pavés, $\mathbf{R}^d$ est muni d'une topologie et donc d'une famille de parties distinguées, les ouverts. On montre alors que l'on peut étendre le volume à la tribu engendrée par les ouverts, dite *tribu borélienne.* Ce schéma fonctionne en fait dans une vaste classe d'espaces topologiques et l'on y dispose alors des théorèmes qui vont nous intéresser ici : les espaces de fonctions intégrables forment des espaces de Banach et les fonctions continues y forment un sous-espace dense.

Étant donné l'approche choisie dans ce livre, nous allons suivre un chemin inverse à celui évoqué ci-dessus : nous allons partir des fonctions continues, pour lesquelles nous disposons déjà d'une théorie satisfaisante de l'intégration, pour définir la classe des fonctions mesurables, qui sont les fonctions les plus générales susceptibles d'être intégrées, puis les parties mesurables et enfin la mesure de Lebesgue.

Définition 11.1 — *Soient* Ω *un ouvert de* $\mathbf{R}^d$ *et* $f : \Omega \longrightarrow \mathbf{R}$ *une fonction numérique. On dit que* f *est mesurable (sur* Ω*) s'il existe une suite* $(f_k)_{k\in\mathbf{N}}$ *de fonctions continues sur* Ω *telle que pour presque tout* $\mathbf{x} \in \Omega$ *on ait* $f(\mathbf{x}) = \lim_{k\to+\infty} f_k(\mathbf{x})$.

Dit de manière courte, une fonction mesurable est limite simple presque partout d'une suite de fonctions continues. Une fonction mesurable est finie presque partout, et toute fonction continue est mesurable.

Exemples :

– Soit $f(x) = \mathrm{Log}\,|x|$ si $x \neq 0$ et $f(0) = 0$ (ou n'importe quelle valeur). Alors f est mesurable sur $\mathbf{R}$, car elle est limite presque partout (pour tout $x \in \mathbf{R}^*$) de la suite de fonctions continues $f_k(x) = \max\{\mathrm{Log}\,|x|, -k\}$.

– Soit $\chi : \mathbf{R} \longrightarrow \mathbf{R}$ la fonction de Dirichlet, indicatrice des nombres rationnels. Alors χ est mesurable, car, étant nulle presque partout, elle est limite de la suite de fonctions nulles $f_k = 0$.

– Si f est mesurable et si $g = f$ presque partout, alors g est mesurable. En particulier, pour vérifier qu'une fonction est mesurable, il suffit de connaître ses valeurs presque partout.

– Soit P un pavé de $\mathbf{R}^d$, alors $\mathbf{1}_\mathrm{P}$ est mesurable. Il en est de même de la fonction indicatrice d'un ouvert, d'un fermé, etc. Nous laissons la démonstration en exercice instructif pour le lecteur. Une conséquence est que les fonctions en escalier (en n'importe quel sens raisonnable) sont mesurables. Une autre est que f est mesurable sur Ω si et seulement si son prolongé par zéro à $\mathbf{R}^d$ tout entier est mesurable.

Avant d'énoncer quelques propriétés générales des fonctions mesurables, nous allons donner deux théorèmes qui prouvent que la mesurabilité est, en quelque sorte, la notion de régularité adaptée à la théorie de l'intégration.

Théorème 11.1 — *Si $f : \Omega \longrightarrow \mathbf{R}$ est mesurable et s'il existe une fonction g Lebesgue-intégrable sur Ω telle que $|f(\mathbf{x})| \leqslant g(\mathbf{x})$ presque partout, alors f est Lebesgue-intégrable sur Ω.*

Démonstration — Commençons par prolonger f et g par zéro à $\mathbf{R}^d$ tout entier ; soit $f_k : \mathbf{R}^d \longrightarrow \mathbf{R}$ une suite de fonctions continues telles que $f(\mathbf{x}) = \lim f_k(\mathbf{x})$ presque partout. On pose $\tilde{f}_k(\mathbf{x}) = \text{méd}\{\mathbf{1}_{[-k,k]^d}(\mathbf{x})f_k(\mathbf{x}), g(\mathbf{x}), -g(\mathbf{x})\}$, où $\text{méd}\{a, b, c\}$ désigne celui des trois nombres réels a, b, c qui est entre les deux autres. Alors $\tilde{f}_k$ est Lebesgue-intégrable sur $\mathbf{R}^d$, vérifie $|\tilde{f}_k(\mathbf{x})| \leqslant |g(\mathbf{x})|$ et $\lim_{k\to+\infty} \tilde{f}_k(\mathbf{x}) = f(\mathbf{x})$ presque partout (car $-g \leqslant f \leqslant g$ p.p.). On peut donc appliquer le théorème de convergence dominée, et f est Lebesgue-intégrable. □

Théorème 11.2 — *Si $f : \Omega \longrightarrow \mathbf{R}$ est Lebesgue-intégrable sur $\mathbf{R}^d$, alors f est mesurable.*

Démonstration — La preuve directe de ce théorème est assez difficile. L'utilisation du théorème de différentiation de Lebesgue 8.16 la simplifie considérablement. Nous allons nous restreindre au cas où f est définie sur un intervalle compact $[a, b]$ de $\mathbf{R}$. Étendons f par zéro à $\mathbf{R}$ tout entier de manière habituelle, en posant $\tilde{f}(x) = f(x)$ si $x \in [a, b]$ et $\tilde{f}(x) = 0$ si $x \notin [a, b]$. La fonction $\tilde{f}$ est Lebesgue-intégrable sur $\mathbf{R}$, et l'on peut considérer son intégrale indéfinie

$$\mathrm{F}(x) = \int_a^x f(t)\,\mathrm{d}t.$$

D'après le théorème 8.3 la fonction F est continue sur $\mathbf{R}$. Enfin, d'après le théorème de différentiation de Lebesgue, elle est dérivable presque partout, de dérivée $\mathrm{F}'(x) = \tilde{f}(x)$.

Mais alors la suite de fonctions $f_k : \mathbf{R} \longrightarrow \mathbf{R}$ définies par

$$f_k(x) = \frac{\mathrm{F}(x+1/k) - \mathrm{F}(x)}{1/k}$$

est une suite de fonctions continues convergeant presque partout vers $\tilde{f}$. Ainsi $\tilde{f}$, et donc f, est mesurable. □

Pour conclure ce chapitre, rassemblons en un unique théorème les propriétés de l'espace $\mathcal{M}(\Omega)$ des fonctions mesurables sur l'ouvert Ω, et qui dit que cette algèbre est stable par toutes les opérations usuelles.

Théorème 11.3 — *Soient $f, g, h \in \mathcal{M}(\Omega)$ ainsi que $\lambda \in \mathbf{R}$.*

1. *Les fonctions λf, $f+g$, $f-g$ et fg sont dans $\mathcal{M}(\Omega)$;*
2. *Si $f(\mathbf{x}) \neq 0$ presque partout, alors $1/f \in \mathcal{M}(\Omega)$;*
3. *Si $k(\mathbf{x}) = f(\mathbf{x})$ presque partout, alors $k \in \mathcal{M}(\Omega)$;*
4. *Si $\varphi : \mathbf{R} \longrightarrow \mathbf{R}$ est continue, alors $\varphi \circ f \in \mathcal{M}(\mathbf{R}^d)$;*
5. *La fonction f est mesurable si et seulement si f^+ et f^- le sont ;*
6. *Les fonctions f et g sont mesurables si et seulement si $\max\{f, g\}$ et $\min\{f, g\}$ le sont ;*
7. *Si f, g, h sont mesurables, alors $\text{méd}\{f, g, h\}$ l'est ;*
8. *La limite simple presque partout d'une suite de fonctions mesurables est mesurable.*

Nous laissons le lecteur écrire une démonstration de ce théorème. C'est un bon exercice pour se familiariser avec la notion de fonction mesurable et il ne présente aucune difficulté.

11.2. Ensembles mesurables, mesure de Lebesgue

En analyse, on est vite amené a intégrer des fonctions sur des parties de $\mathbf{R}$ ou $\mathbf{R}^d$ qui ne sont pas des intervalles ou des pavés, ni même des ouverts, mais des objets obtenus par la prise, par exemple, d'intersection ou d'union de familles dénombrables. Considérons par exemple une suite de fonctions f_n définies sur $\mathbf{R}$, et supposons que l'on ne s'intéresse qu'aux points x de $\mathbf{R}$ pour lesquels $f_n(x)$ tend vers 0. Cela s'écrit, avec les quantificateurs :

$$\forall \varepsilon > 0,\ \exists \mathrm{N} \in \mathbf{N} \ \text{ t.q. } \ \forall n \geqslant \mathrm{N}, \quad |f_n(x)| < \varepsilon.$$

Il se trouve que dans $\mathbf{R}$, être arbitrairement petit peut s'écrire uniquement avec des entiers, et l'on peut donc écrire la condition ci-dessus sous la forme suivante :

$$\forall \mathrm{K} \in \mathbf{N}^*,\ \exists \mathrm{N} \in \mathbf{N} \ \text{ t.q. } \ \forall n \geqslant \mathrm{N}, \quad |f_n(x)| < \mathrm{K}^{-1}.$$

Nous pouvons maintenant décrire l'ensemble des x tels que $f_n(x)$ tend vers 0 en remplaçant les signes $\forall$ par une intersection et les signes $\exists$ par une union, et l'on découvre la relation suivante :

$$\left\{x \in \mathbf{R},\ \lim_{n\to+\infty} f_n(x) = 0\right\} = \bigcap_{\mathrm{K}\in\mathbf{N}^*} \bigcup_{\mathrm{N}\in\mathbf{N}} \bigcap_{n\geqslant \mathrm{N}} \left\{x \in \mathbf{R},\ |f_n(x)| < \frac{1}{\mathrm{K}}\right\}.$$

Si les fonctions f_n sont continues, on obtient ainsi une intersection dénombrable d'une union dénombrable d'intersections dénombrables (ouf !) d'ouverts de $\mathbf{R}$. Si l'on souhaite intégrer sur une telle partie E, on aimerait en particulier pouvoir contrôler sa taille, de manière à pouvoir écrire des inégalités du genre

$$\int_{\mathrm{E}} |f(x)|\,\mathrm{d}x \leqslant |\mathrm{E}| \sup\left\{|f(x)|, x \in \mathrm{E}\right\},$$

où $|\mathrm{E}|$ désignerait la taille recherchée de E. Si E est un intervalle $[a, b]$, sa taille est naturellement sa longueur $|b-a|$. De même, si E est un ouvert, on sait qu'il est réunion dénombrable d'intervalles ouverts disjoints de $\mathbf{R}$, et sa taille devrait être la somme des longueurs de ces intervalles. Mais que dire d'une partie E comme ci-dessus ? Nous allons donner une réponse à cette question, en montrant que la notion de taille, que nous appellerons dorénavant la *mesure,* s'étend à ce genre de sous-ensemble compliqué de $\mathbf{R}$ ou $\mathbf{R}^d$, grâce à la notion de fonction mesurable et à l'intégrale de Lebesgue.

Si P est un pavé compact de $\mathbf{R}^d$, il semble naturel de dire qu'il est mesurable et que sa mesure est simplement son volume $|\mathrm{P}|$. Plus généralement, on pose la définition suivante :

Définition 11.2 — *On dit qu'une partie* E *de* $\mathbf{R}^d$ *est mesurable si sa fonction caractéristique* $\mathbf{1}_{\mathrm{E}}$ *est une fonction mesurable. Si de plus* $\mathbf{1}_{\mathrm{E}}$ *est intégrable, on dit que* E *est de mesure finie et la quantité*

$$|\mathrm{E}| = \int_{\mathbf{R}^d} \mathbf{1}_{\mathrm{E}}(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

s'appelle la mesure de Lebesgue de E.

Vous pouvez, à titre d'exercice, essayer de démontrer que l'ensemble E ci-dessus, quand les fonctions f_n sont toutes continues, est une partie mesurable de $\mathbf{R}$. La notion de mesurabilité est, comme pour les fonctions, une notion de régularité : une partie est mesurable si elle n'est pas trop atroce. Cette notion d'atrocité est cependant très lâche, et il est très difficile de démontrer l'existence d'ensembles non mesurables. La proposition suivante montre que pour un ensemble mesurable, le problème de la finitude de sa mesure ne se pose que s'il est non borné :

Proposition 11.4 — *Si* E *est un ensemble mesurable borné, il est de mesure finie.*

Démonstration — Soit en effet P un pavé compact contenant E. Alors $\mathbf{1}_{\mathrm{E}}$ est mesurable, tandis que $\mathbf{1}_{\mathrm{P}}$ est Lebesgue-intégrable, et l'on a l'inégalité $\mathbf{1}_{\mathrm{E}} \leqslant \mathbf{1}_{\mathrm{P}}$. D'après le théorème 11.1, la fonction $\mathbf{1}_{\mathrm{E}}$ est Lebesgue-intégrable, et donc E est de mesure finie. □

Avant de passer aux propriétés générales des ensembles mesurables et de la mesure de Lebesgue, levons une ambiguïté terminologique que le lecteur attentif n'aura pas manqué de noter : les ensembles de mesure nulle (au sens de la définition vue dans un des chapitres précédents) sont-ils aussi les ensembles mesurables de mesure de Lebesgue nulle ? La réponse est, fort heureusement, oui.

Proposition 11.5 — *Une partie* E *de* $\mathbf{R}^d$ *est de mesure nulle si et seulement si elle est mesurable, de mesure de Lebesgue nulle.*

Démonstration — Si E est de mesure nulle, on sait que $\mathbf{1}_{\mathrm{E}}$ est une fonction positive nulle presque partout, et donc elle est Lebesgue-intégrable d'intégrale nulle. Elle est en particulier mesurable, donc E est mesurable et E est de mesure de Lebesgue nulle puisque

$$\int_{\mathbf{R}^d} \mathbf{1}_{\mathrm{E}} = 0.$$

Si réciproquement E est mesurable de mesure de Lebesgue nulle, alors $\mathbf{1}_{\mathrm{E}}$ est positive, intégrable et d'intégrale nulle. Elle est donc nulle presque partout et donc E est de mesure nulle. □

Pour l'instant, nous ne disposons que de peu d'exemples de parties mesurables : les pavés et les ensembles de mesure nulle. Cependant l'utilisation des théorèmes de convergence monotone et dominée permet de démontrer que la classe des ensembles mesurables est stable par de nombreuses opérations. Avant de donner un énoncé général, illustrons notre propos en montrant que tout ouvert borné de $\mathbf{R}^d$ est mesurable de mesure de Lebesgue finie.

Soit Ω un ouvert borné de $\mathbf{R}^d$ et soit P un pavé compact contenant Ω. Pour tout entier positif k et tout d-uplet $(x_1, \ldots, x_d) \in \mathbf{Z}^d$, posons

$$\mathrm{P}_k(x_1, \ldots, x_d) = \left[\frac{x_1}{2^k}, \frac{x_1+1}{2^k}\right[\times \cdots \times \left[\frac{x_d}{2^k}, \frac{x_d+1}{2^k}\right[.$$

Si k est fixé, la famille des pavés semi-ouverts P_k, quand $(x_1, \ldots, x_d)$ varie dans $\mathbf{Z}^d$ forme une partition de $\mathbf{R}^d$ (en atomes de plus en plus petits à mesure que k grandit). Notons

$$\mathrm{S}_\Omega^k = \{(x_1, \ldots, x_d) \in \mathbf{Z}^d, \mathrm{P}_k(x_1, \ldots, x_d) \subset \Omega\}.$$

Comme Ω est ouvert, pour tout $\mathbf{x} \in \Omega$ il existe un entier $k \geqslant 0$ tel que $\mathrm{B}(\mathbf{x}, 2^{-k}) \subset \Omega$ et l'on peut affirmer que

$$\Omega = \bigcup_{k \geqslant 0} \bigcup_{(x_1, \ldots, x_d) \in \mathrm{S}_\Omega^k} \mathrm{P}_k(x_1, \ldots, x_d).$$

Si l'on pose

$$\varphi_k = \sum_{(x_1, \ldots, x_d) \in \mathrm{S}_\Omega^k} \mathbf{1}_{\mathrm{P}_k(x_1, \ldots, x_d)},$$

$\varphi_k(\mathbf{x})$ tend vers $\mathbf{1}_\Omega(\mathbf{x})$ pour tout $\mathbf{x} \in \mathbf{R}^d$. La fonction φ_k étant en escalier, est Lebesgue-intégrable. Enfin, la suite $(\varphi_k(\mathbf{x}))$ est croissante, pour tout $\mathbf{x} \in \mathbf{R}^d$ on a $0 \leqslant \varphi_k(\mathbf{x}) \leqslant \mathbf{1}_\Omega(\mathbf{x}) \leqslant \mathbf{1}_\mathrm{P}(\mathbf{x})$ et, comme $\mathbf{1}_\mathrm{P}$ est Lebesgue-intégrable, on peut appliquer à la suite φ_k le théorème de convergence monotone : $\mathbf{1}_\Omega$ est donc Lebesgue-intégrable, et ainsi Ω est mesurable de mesure finie.

Nous allons maintenant énoncer une liste de propriétés vérifiées par les ensembles mesurables. Afin de ne pas avoir à distinguer entre les ensembles de mesure finie et les autres, nous poserons que $|\mathrm{E}| = +\infty$ si E est mesurable mais pas de mesure finie. Nous notons $\mathbf{M}$ l'ensemble des parties mesurables de $\mathbf{R}^d$.

1. L'ensemble vide est élément de $\mathbf{M}$, et $|\emptyset| = 0$;
2. L'ensemble $\mathbf{R}^d$ en entier est mesurable, de mesure infinie ;
3. Tout pavé (produit d'intervalles de $\mathbf{R}$) est mesurable, et sa mesure est égale à son volume ;
4. Si A, B sont deux éléments de $\mathbf{M}$, alors
   – $\mathrm{A} \cap \mathrm{B}$ et $\mathrm{A} \cup \mathrm{B}$ sont éléments de $\mathbf{M}$ et $|\mathrm{A} \cup \mathrm{B}| + |\mathrm{A} \cap \mathrm{B}| = |\mathrm{A}| + |\mathrm{B}|$,
   – $\mathrm{A} - \mathrm{B}$ est élément de M ; en particulier, $\mathbf{M}$ est stable par passage au complémentaire,
   – si $\mathrm{B} \subset \mathrm{A}$ et si B est de mesure finie, alors $|\mathrm{A} - \mathrm{B}| = |\mathrm{A}| - |\mathrm{B}|$ et en particulier $|\mathrm{B}| \leqslant |\mathrm{A}|$;
5. Si $\mathrm{B}_1 \supseteq \mathrm{B}_2 \supseteq \cdots \supseteq \mathrm{A}_n \supseteq \cdots$ est une suite décroissante d'éléments de $\mathbf{M}$, alors leur intersection est élément de $\mathbf{M}$ et si l'un des B_n est de mesure finie alors
$$\left| \bigcap_{n \geqslant 1} \mathrm{B}_n \right| = \lim_{n \to +\infty} |\mathrm{B}_n| = \inf_{n \geqslant 1} |\mathrm{B}_n| \, ;$$
6. Si $\mathrm{A}_1 \subseteq \mathrm{A}_2 \subseteq \cdots \subseteq \mathrm{B}_n \subseteq \cdots$ est une suite croissante d'éléments de $\mathbf{M}$, alors leur union est élément de $\mathbf{M}$ et
$$\left| \bigcup_{n \geqslant 1} \mathrm{A}_n \right| = \lim_{n \to +\infty} |\mathrm{A}_n| = \sup_{n \geqslant 1} |\mathrm{A}_n| \, ;$$
7. Si $\{\mathrm{C}_i, i \in \mathbf{N}\}$ est une famille dénombrable d'éléments de $\mathbf{M}$ deux à deux disjoints, alors leur union est mesurable et
$$\left| \bigcup_{i \in \mathbf{N}} \mathrm{C}_i \right| = \sum_{i \in \mathbf{N}} |\mathrm{C}_i| \, ;$$

8. Si $\{D_i, i \in \mathbf{N}\}$ est une famille dénombrable d'éléments de $\mathbf{M}$, alors l'union et l'intersection des D_i est dans $\mathbf{M}$ et on a
$$\left|\bigcup_{n\in\mathbf{N}} D_n\right| \leqslant \sum_{n\in\mathbf{N}} |D_n|.$$

La démonstration de toutes ces propriétés est simple, en utilisant les définitions de fonction et partie mesurable, puis les théorèmes de convergence monotone et dominée.

Une famille de parties de $\mathbf{R}^d$ (ou, en général, d'un ensemble quelconque) qui contient l'ensemble vide, l'ensemble total, est stable par passage au complémentaire et par intersection dénombrable s'appelle une *σ-algèbre,* ou une *tribu.* La famille $\mathbf{M}$ est donc une tribu de parties de $\mathbf{R}$, qui contient la famille des ouverts et des fermés. La plus petite tribu de parties de $\mathbf{R}^d$ contenant les ouverts s'appelle la *tribu borélienne,* notée $\mathrm{Bor}(\mathbf{R}^d)$. Un de ses éléments s'appelle un borélien. Un borélien est donc mesurable, mais la réciproque est fausse. Il y a beaucoup plus de mesurables que de boréliens, mais l'on peut dans de nombreux problèmes d'analyse se restreindre à la tribu borélienne. La mesure de Lebesgue définit sur les tribus $\mathbf{M}$ et $\mathrm{Bor}(\mathbf{R}^d)$ ce que l'on appelle une *mesure σ-additive,* le σ référant au bon comportement sous les opérations dénombrables. Le langage des tribus et des mesure est le langage naturel de la théorie des *probabilités.*

Nous verrons dans la suite toute la puissance des propriétés énoncées ci-dessus concernant les parties mesurables. Nous en donnons immédiatement un corollaire très utilisé, en particulier en théorie des probabilités. Si (E_n) est une famille dénombrable de parties mesurables de $\mathbf{R}^d$, on note respectivement
$$E_\star = \bigcup_{n=1}^{+\infty} \bigcap_{k\geqslant n} E_k$$
et
$$E^\star = \bigcap_{n=1}^{+\infty} \bigcup_{k\geqslant n} E_k\,;$$
sous leurs apparences compliquées, ces deux parties ont une interprétation simple : $\mathbf{x}$ est élément de $E_\star$ si et seulement s'il appartient à tous les E_n sauf un nombre fini d'entre eux ; de même $\mathbf{x}$ est élément de $E^\star$ si et seulement si $\mathbf{x}$ appartient à une infinité de E_n. D'après ce qui précède, $E_\star$ et $E^\star$ sont mesurables. Mais on a mieux : pour tout $\mathbf{x}$, on a
$$\mathbf{1}_{E_\star}(\mathbf{x}) = \liminf_{n\to+\infty} \mathbf{1}_{E_n}(\mathbf{x}) \quad \text{et} \quad \mathbf{1}_{E^\star}(\mathbf{x}) = \limsup_{n\to+\infty} \mathbf{1}_{E_n}(\mathbf{x}).$$

On déduit du lemme de Fatou que s'il existe une constante M telle que $|E_n| \leqslant M$ pour tout n, alors on a les inégalités
$$|E_\star| \leqslant \liminf |E_n| \leqslant \limsup |E_n| \leqslant |E^\star| \leqslant M.$$

Lemme de Borel-Cantelli — *Soient* P *un pavé compact de* $\mathbf{R}^d$ *et* (E_n) *une famille dénombrable de parties mesurables de* P. *Si la série* $\sum |\mathrm{E}_n|$ *converge, alors* $|\mathrm{E}_\star| = |\mathrm{E}^\star| = 0$.

Démonstration — D'après ce qui précède, $0 \leqslant |\mathrm{E}_\star| \leqslant |\mathrm{E}^\star| \leqslant |\mathrm{P}|$, et il nous suffit donc de démontrer que $|\mathrm{E}^\star| = 0$. Mais pour tout $n \geqslant 0$ on a

$$\mathrm{E}^\star \subseteq \bigcup_{k \geqslant n} \mathrm{E}_k$$

et donc, d'après les propriétés de la mesure de Lebesgue et des mesurables énoncées plus haut,

$$|\mathrm{E}^\star| \leqslant \left| \bigcup_{k \geqslant n} \mathrm{E}_k \right| \leqslant \sum_{k=n}^{+\infty} |\mathrm{E}_k|.$$

Le dernier terme est le reste d'ordre n d'une série convergente, il tend donc vers 0 quand n tend vers l'infini, et le lemme de Borel-Cantelli est démontré. □

11.3. Intégrale sur des parties mesurables

Les notions de mesure (de Lebesgue, ici), d'ensemble et de fonction mesurable permettent de considérer l'intégration de fonctions définies sur des parties plus générales que des pavés ou des ouverts. Soit en effet f une fonction Lebesgue-intégrable sur $\mathbf{R}^d$ et E une partie mesurable. Alors f et $\mathbf{1}_\mathrm{E}$ sont mesurables et donc leur produit $\mathbf{1}_\mathrm{E} f$ l'est aussi. Mais $\mathbf{1}_\mathrm{E} f \leqslant f$ et donc $\mathbf{1}_\mathrm{E} f$ est Lebesgue-intégrable, et on appelle son intégrale l'intégrale de f sur la partie E, ce que l'on note

$$\int_\mathrm{E} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbf{R}^d} \mathbf{1}_\mathrm{E}(\mathbf{x}) f(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Réciproquement, si f n'est définie que sur E, on la prolonge par zéro à $\mathbf{R}^d$ tout entier, et on dit qu'elle est Lebesgue-intégrable sur E si sa fonction prolongée est Lebesgue-intégrable sur $\mathbf{R}^d$; on utilise alors les mêmes notations.

Cette possibilité d'intégrer sur des parties mesurables quelconques est très agréable et possède toutes les propriétés que l'on peut intuitivement désirer ; ainsi si f est intégrable sur deux parties disjointes E_1 et E_2, alors f est intégrable sur $\mathrm{E}_1 \cup \mathrm{E}_2$ et l'intégrale sur l'union est la somme des intégrales (une forme générale de la relation de Chasles) ; de même, si f est majorée par une constante M, alors

$$\int_\mathrm{E} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \leqslant |\mathrm{E}|\mathrm{M}.$$

Nous allons immédiatement utiliser cette possibilité en démontrant un théorème qui va nous être très utile pour la suite, l'inégalité de Tchebicheff.

Théorème 11.6 — *Soient f une fonction Lebesgue-intégrable sur $\mathbf{R}^d$ et $r > 0$. Alors la partie $\mathrm{K}_r = \{\mathbf{x} \in \mathbf{R}^d, f(\mathbf{x}) \geqslant r\}$ est mesurable et l'on a*

$$|\mathrm{K}_r| \leqslant \frac{1}{r}\int_{\mathrm{K}_r} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x} \leqslant \frac{1}{r}\|f\|_{\mathrm{L}^1},$$

où, par définition, on a posé

$$\|f\|_{\mathrm{L}^1} = \int_{\mathbf{R}^d} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x},$$

qui s'appelle la norme L^1 de f.

Démonstration — Remarquons tout d'abord qu'en remplaçant f par f^+, on peut supposer que f est positive. Enfin, pour démontrer que K_r est mesurable, il suffit de prouver que son complémentaire l'est :

$$\mathbf{R}^d - \mathrm{K}_r = \big\{\mathbf{x} \in \mathbf{R}^d, 0 \leqslant f(\mathbf{x}) < r\big\}.$$

Pour tout entier $n \geqslant 0$ et tout réel $\mathrm{R} > 0$, posons

$$\varphi_n(\mathbf{x}) = \min\big\{1, n\max\{r - f(\mathbf{x}), 0\}\big\}\,.\,\mathbf{1}_{[-\mathrm{R},\mathrm{R}]^d}(\mathbf{x}).$$

La fonction φ_n est mesurable et vérifie $0 \leqslant \varphi_n \leqslant 1$; comme de plus φ_n est nulle hors du pavé compact $[-\mathrm{R},\mathrm{R}]^d$, elle est Lebesgue-intégrable. Enfin, la suite φ_n est croissante et converge simplement vers la fonction indicatrice de $(\mathbf{R}^d - \mathrm{K}_r) \cap [-\mathrm{R},\mathrm{R}]^d$. On peut donc appliquer le théorème de convergence monotone et la fonction indicatrice de $(\mathbf{R}^d - \mathrm{K}_r) \cap [-\mathrm{R},\mathrm{R}]^d$ est intégrable, donc mesurable. En écrivant

$$\mathbf{R}^d - \mathrm{K}_r = \bigcup_{\mathrm{N}\in\mathbf{N}} (\mathbf{R}^d - \mathrm{K}_r) \cap [-\mathrm{N},\mathrm{N}]^d,$$

on en déduit que $\mathbf{R}^d - \mathrm{K}_r$ est mesurable.

Pour achever la preuve de l'inégalité de Tchebicheff, notons que

$$r\mathbf{1}_{\mathrm{K}_r}(\mathbf{x}) \leqslant \mathbf{1}_{\mathrm{K}_r} f(\mathbf{x})$$

et donc

$$r|\mathrm{K}_r| \leqslant \int_{\mathrm{K}_r} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x} \leqslant \int_{\mathbf{R}^d} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x},$$

ce qui démontre le théorème. □

Remarquons que dans la preuve du théorème, nous avons quasiment démontré la proposition suivante :

Proposition 11.7 — *Une fonction $f : \mathbf{R}^d \longrightarrow \mathbf{R}$ est mesurable si et seulement si pour tout intervalle I de $\mathbf{R}$, $f^{-1}(\mathrm{I})$ est une partie mesurable de $\mathbf{R}^d$.*

Nous laissons le lecteur en écrire une démonstration complète.

11.4. La convergence monotone revisitée

Nous allons utiliser l'inégalité de Tchebicheff pour démontrer une version encore plus souple du théorème de convergence monotone, où l'on n'aura même plus à faire l'hypothèse que la suite de fonctions converge presque partout vers une limite finie. Cette version nous permettra ensuite de démontrer le théorème de Beppo Levi, ingrédient crucial pour la preuve de la complétude des espaces de Lebesgue L^p.

Théorème 11.8 — *Soit (φ_n) une suite de fonctions Lebesgue-intégrables sur $\mathbf{R}^d$, telle que la suite $(\varphi_n(x))$ soit croissante presque partout et que la suite des intégrales*

$$\int_{\mathbf{R}^d} \varphi_n$$

soit majorée. Alors

1. L'ensemble

$$\mathrm{Z} = \left\{\mathbf{x} \in \mathbf{R}^d, \lim_{n\to+\infty} \varphi_n(\mathbf{x}) = +\infty\right\}$$

est de mesure nulle ;

2. Si l'on pose $\varphi(\mathbf{x}) = \lim \varphi_n(\mathbf{x})$ pour tout $\mathbf{x}$ hors de Z et $\varphi(\mathbf{x}) = 0$ pour $\mathbf{x} \in \mathrm{Z}$, alors φ est Lebesgue-intégrable et l'on a les deux identités

$$\int_{\mathbf{R}^d} \varphi(\mathbf{x})\,\mathrm{d}\mathbf{x} = \lim_{n\to+\infty} \int_{\mathbf{R}^d} \varphi_n(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

et

$$\lim_{n\to+\infty} \int_{\mathbf{R}^d} |\varphi_n(\mathbf{x}) - \varphi(\mathbf{x})|\,\mathrm{d}\mathbf{x} = 0.$$

Démonstration — Soit pour commencer $\tilde\varphi : \mathbf{R}^d \longrightarrow \mathbf{R} \cup \{+\infty\}$ la limite simple de la suite φ_n ; $\tilde\varphi(\mathbf{x}) = \varphi(\mathbf{x})$ si $\mathbf{x} \notin \mathrm{Z}$ et $\tilde\varphi(\mathbf{x}) = +\infty$ si $\mathbf{x} \in \mathrm{Z}$. Pour deux entiers $m, n \geqslant 1$ on pose

$$\mathrm{E}_{n,m} = \left\{\mathbf{x} \in \mathbf{R}^d, \varphi_n(\mathbf{x}) \geqslant m\right\}.$$

Les parties $\mathrm{E}_{n,m}$ vérifient la condition de décroissance $\mathrm{E}_{n,m} \subseteq \mathrm{E}_{n+1,m}$ et d'après l'inégalité de Tchebicheff elles sont mesurables de mesure satisfaisant $|\mathrm{E}_{n,m}| \leqslant \mathrm{K}m^{-1}$, où $\mathrm{K} = \sup_n \|\varphi_n\|_{\mathrm{L}^1}$. Mais on peut exprimer Z en fonction des $\mathrm{E}_{n,m}$:

$$\mathrm{Z} = \bigcap_{m\geqslant 1} \bigcup_{\mathrm{N}\geqslant 1} \bigcap_{n\geqslant \mathrm{N}} \mathrm{E}_{n,m}.$$

La partie Z est donc mesurable et comme, par la condition de décroissance, on a

$$\left| \bigcup_{\mathrm{N}\geqslant 0} \bigcap_{n\geqslant \mathrm{N}} \mathrm{E}_{n,m} \right| \leqslant \frac{\mathrm{K}}{m}$$

et donc $|\mathrm{Z}| = 0$.

On peut maintenant conclure, par le théorème de convergence dominée usuel, que φ est Lebesgue-intégrable et que

$$\int_{\mathbf{R}^d} \varphi(\mathbf{x})\,\mathrm{d}\mathbf{x} = \lim_{n\to+\infty} \varphi_n(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Par ailleurs, par monotonicité de la suite φ_n on a $|\varphi(\mathbf{x}) - \varphi_n(\mathbf{x})| \leqslant |\varphi(\mathbf{x}) - \varphi_0(\mathbf{x})|$ presque partout et d'autre par $|\varphi(\mathbf{x}) - \varphi_n(\mathbf{x})|$ tend vers 0 presque partout. On peut donc appliquer le théorème de convergence dominée, et en déduire que

$$\lim_{n\to+\infty} \int_{\mathbf{R}^d} |\varphi_n(\mathbf{x}) - \varphi(\mathbf{x})|\,\mathrm{d}\mathbf{x} = \lim_{n\to+\infty} \|\varphi_n - \varphi\|_{\mathrm{L}^1} = 0. \qquad \square$$

Nous pouvons maintenant énoncer le théorème de Beppo Levi. Nous allons adopter une approche consistant à non plus parler de fonctions Lebesgue-intégrables, mais de points (ou vecteurs) de l'espace $\mathcal{L}^1$, afin de nous préparer à l'étude des espaces fonctionnels en tant que tels.

Théorème de Beppo Levi — *Soit* (ψ_k) *une suite d'éléments de l'espace* $\mathcal{L}^1(\mathbf{R}^d)$ *telle que la série*

$$\sum_{k=0}^{+\infty} \|\psi_k\|_{\mathrm{L}^1}$$

soit convergente, de somme K. *Alors la série*

$$\sum_{k=0}^{+\infty} \psi_k(\mathbf{x})$$

converge absolument presque partout, la fonction somme est une fonction Lebesgue-intégrable et

$$\int_{\mathbf{R}^d} \left(\sum_{k=0}^{+\infty} \psi_k(\mathbf{x}) \right) \mathrm{d}\mathbf{x} = \sum_{k=0}^{+\infty} \int_{\mathbf{R}^d} \psi_k(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Démonstration — On pose

$$\tau_n(\mathbf{x}) = \sum_{k=0}^{n} \psi_k(\mathbf{x}) \text{ et } \varphi_n(\mathbf{x}) = \sum_{k=0}^{n} |\psi_k(\mathbf{x})|.$$

Le théorème de convergence monotone revisité implique l'existence d'une partie de mesure nulle Z de $\mathbf{R}^d$ telle que pour tout $\mathbf{x} \notin \mathrm{Z}$ on ait $\lim \varphi_n(\mathbf{x}) = \varphi(\mathbf{x})$, où φ est un élément de $\mathcal{L}^1(\mathbf{R}^d)$.

Pour tout $\mathbf{x} \notin \mathrm{Z}$, la série $\sum \psi_k(\mathbf{x})$ est une série absolument convergente de nombres réels et donc elle est convergente. En particulier

$$\tau(\mathbf{x}) = \sum_{k=0}^{+\infty} \psi_k(\mathbf{x})$$

existe et est finie pour presque tout $\mathbf{x}$. Comme par ailleurs $|\tau_k(\mathbf{x})| \leqslant \varphi(\mathbf{x})$ pour presque tout $\mathbf{x}$, le théorème de convergence dominée permet d'achever la démonstration du théorème. □

11.5. Complétude des espaces L^p

Nous allons maintenant commencer l'étude de l'espace des fonctions Lebesgue-intégrables muni de sa norme naturelle, la « norme L^1 ». Nous l'avons déjà introduite plus haut mais vu son importance nous redonnons sa définition :

Définition 11.3 — *La norme* L^1 *d'une fonction Lebesgue-intégrable sur* $\mathbf{R}^d$ *est par définition le nombre réel*

$$\|f\|_{\mathrm{L}^1} = \int_{\mathbf{R}^d} |f(\mathbf{x})| \, \mathrm{d}\mathbf{x}.$$

Bien sûr, si l'on s'intéresse aux fonctions Lebesgue-intégrables sur une partie (mesurable) de $\mathbf{R}^d$, on modifie en conséquence la définition. Nous allons voir tout de suite que cette définition est quelque peu abusive au sens où la norme L^1 n'est pas tout-à-fait une norme, mais l'usage de cette terminologie sera justifiée un peu plus bas.

Rappelons que l'on a démontré que l'espace $\mathcal{L}^1(\mathbf{R}^d)$ est un $\mathbf{R}$-espace vectoriel. Sur un $\mathbf{R}$-espace vectoriel V, une application $\|\cdot\| : \mathrm{V} \longrightarrow \mathbf{R}$ est une *semi-norme* si elle vérifie les propriétés suivantes :

(P) $\qquad \forall x \in \mathrm{V}, \quad \|x\| \geqslant 0$

(S′) $\qquad x = 0 \implies \|x\| = 0$

(H) $\qquad \forall x \in \mathrm{V}, \forall \lambda \in \mathbf{R}, \quad \|\lambda x\| = |\lambda| \|x\|$

(T) $\qquad \forall x, y \in \mathrm{V}, \quad \|x + y\| \leqslant \|x\| + \|y\|.$

La condition (T) s'appelle l'inégalité triangulaire. Si dans la condition (S′) on a une équivalence :

(S) $\qquad \|x\| = 0 \iff x = 0$

on dit que la semi-norme est une *norme* tout court. Muni d'une semi-norme, l'espace V peut être l'objet d'une étude topologique ou géométrique : on peut y définir les boules, ouvertes ou fermées, la (pseudo) distance entre deux éléments ($d(x, y) = \|x - y\|$), la

notion de partie ouverte, fermée, bornée, etc. Si la semi-norme est une norme, l'espace est *séparé,* au sens où si deux éléments sont à distance nulle l'un de l'autre, alors ils sont égaux. Cela a pour corollaire que si une suite converge, sa limite est unique.

Lemme 11.9 — *La norme* L^1 *définit une semi-norme sur l'espace* $\mathcal{L}^1(\mathbf{R}^d)$.

La vérification de cette affirmation est facile, et découle de la linéarité et de la positivité de l'intégrale. Notons que ce n'est pas une norme au sens strict du mot, mais que l'on peut juste dire que $\|f\|_{\mathrm{L}^1} = 0$ si et seulement si f est nulle presque partout (et pas nécessairement partout). On peut cependant faire de $\mathcal{L}^1$ un véritable espace normé en décrétant qu'une fonction qui est nulle presque partout est nulle. Cela se concrétise mathématiquement par un passage au quotient.

Définition 11.4 — *Soit* $\mathcal{N} \subset \mathcal{L}^1(\mathbf{R}^d)$ *le sous-espace constitué des fonctions nulles presque partout. L'espace de Lebesgue* $\mathrm{L}^1(\mathbf{R}^d)$ *est par définition l'espace quotient*

$$\mathrm{L}^1(\mathbf{R}^d) = \mathcal{L}^1(\mathbf{R}^d)/\mathcal{N}.$$

Comme quotient d'un espace vectoriel par un de ses sous-espaces, $\mathrm{L}^1(\mathbf{R}^d)$ est un espace vectoriel. Un élément φ de L^1 est une *classe* de fonctions modulo l'égalité presque partout. Si f et g sont dans la classe φ, c'est-à-dire sont des *fonctions* Lebesgue-intégrables telles que $\varphi = [f] = [g]$ (on note ici $[f]$ la classe de f dans L^1), alors $f = g$ presque partout puisque $f - g \in \mathcal{N}$, et l'on peut donc poser $\|\varphi\|_{\mathrm{L}^1} = \|f\|_{\mathrm{L}^1} = \|g\|_{\mathrm{L}^1}$. Cela munit $\mathrm{L}^1(\mathbf{R}^d)$ d'une semi-norme qui, par construction, est une norme, puisque si $\|\varphi\|_{\mathrm{L}^1} = 0$ alors φ est la classe d'une fonction nulle presque partout, c'est-à-dire d'un élément de $\mathcal{N}$.

Pour pouvoir faire de l'analyse ou de la topologie sur un espace muni d'une norme, une condition pratique est de pouvoir démontrer la convergence d'une suite sans connaître *a priori* la valeur de sa limite. Les espaces $\mathbf{R}$ ou $\mathbf{R}^d$, munis de leurs normes usuelles, possèdent cette propriété grâce au critère de Cauchy : on dit qu'ils sont complets.

Définition 11.5 — *On dit qu'une suite* (x_n) *d'éléments d'un espace vectoriel* V *muni d'une semi-norme* $\|\cdot\|$ *est de Cauchy si pour tout* $\varepsilon > 0$ *il existe un* $\mathrm{N} \in \mathbf{N}$ *tel que, pour tous* $n, m \geqslant \mathrm{N}$ *on ait* $\|x_m - x_n\| < \varepsilon$.

On dit que $(\mathrm{V}, \|\cdot\|)$ *est un espace normé complet si toute suite de Cauchy admet une limite.*

Remarquons qu'une suite convergente est nécessairement de Cauchy et donc, dans un espace complet, être de Cauchy est une caractérisation de la convergence. Remarquons également que la définition de complétude est valable pour une semi-norme, mais que la limite d'une suite n'est pas nécessairement unique.

On peut démontrer que si V est de dimension finie, alors il est complet pour n'importe quelle norme dont on le munisse. En dimension infinie, il n'en est plus du tout ainsi. Voici un exemple simple : prenons pour V l'espace des fonctions polynomiales sur l'intervalle $[0, \pi]$, muni de la norme de la convergence uniforme :

$$\|P - Q\| = \sup\big\{|P(x) - Q(x)|,\ x \in [0, 1]\big\}.$$

Considérons la suite P_n d'éléments de V définie par

$$P_n(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots + \frac{x^n}{n!},\ x \in [0, \pi].$$

Cette suite est de Cauchy, car pour tout $n \geqslant m \geqslant 0$ on a

$$\|P_n - P_m\| \leqslant \sum_{k=m+1} n\frac{\pi^n}{n!},$$

et le second membre est arbitrairement petit car la série $\sum \pi^n (n!)^{-1}$ converge (vers e^π). Mais la suite P_n ne converge vers aucun élément P de V. En effet, si P était cette limite, on aurait $P'(x) = P(x)$ pour tout $x \in [0, \pi]$, et cela est impossible si P est un polynôme non nul. En fait, la limite est dans l'espace des fonctions continues sur $[0, \pi]$, et s'y identifie à la fonction exponentielle.

L'espace des fonctions polynomiales est un espace de mauvais aloi pour l'analyse. Il n'est en fait jamais complet, quelle que soit la norme dont on le munisse. Fort heureusement, l'analyste a par ailleurs de nombreux espaces de fonctions à se mettre sous la dent et tous, ou peu s'en faut, sont complets. Le but principal de cette section est de démontrer le théorème suivant (connu sous le nom de Fischer-Riesz) qui, après le théorème de convergence dominée de Lebesgue, est sans doute le théorème le plus important énoncé dans ce traité.

Théorème 11.10 — *L'espace de Lebesgue* $L^1(\mathbf{R}^d)$*, muni de sa norme* L^1*, est un espace normé complet. On dit que c'est un espace de Banach.*

Dans cet espace de classes de fonctions, toute suite de Cauchy converge vers une limite et une seule. Cependant, au niveau des fonctions, la limite n'est unique qu'à l'addition près d'une fonction nulle presque partout. Dans la suite, nous confondrons sans vergogne *classe de fonction* et *fonction.* Cela n'aura aucune incidence sur la justesse des raisonnements car nous nous contenterons d'étudier des *suites* d'éléments de L^1 et que l'union d'une famille *dénombrable* de parties de mesure nulle est elle même de mesure nulle. Voici une illustration de ce phénomène sur lequel nous ne nous étendrons pas plus avant.

Définition 11.6 — *On dit qu'une suite d'éléments de* $\mathcal{L}^1(\mathbf{R}^d)$ *converge en moyenne, ou au sens* L^1*, vers* $f \in \mathcal{L}^1(\mathbf{R}^d)$ *si*

$$\lim_{n\to+\infty} \|f_n - f\|_{L^1} = 0.$$

Si (f_n) converge en moyenne vers f, il est facile de vérifier qu'alors

$$\lim_{n\to+\infty}\int_{\mathbf{R}^d} f_n(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbf{R}^d} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \text{ et } \lim_{n\to+\infty}\|f_n\|_{\mathrm{L}^1} = \|f\|_{\mathrm{L}^1}.$$

Mais on a mieux :

Proposition 11.11 — *Une suite* (φ_n) *d'éléments de* $\mathrm{L}^1(\mathbf{R}^d)$ *converge au sens de la norme* L^1 *vers la classe* $\varphi \in \mathrm{L}^1(\mathbf{R}^d)$ *si et seulement si, pour tout choix de représentants* f_n *de* φ_n *et* f *de* φ *la suite de* fonctions (f_n) *converge en moyenne vers la fonction* f*.*

Nous démontrerons donc le théorème de complétude sous la forme suivante :

Théorème 11.12 — *L'espace* $\mathcal{L}^1(\mathbf{R}^d)$ *muni de la semi-norme* L^1 *est complet.*

Démonstration — Nous allons raisonner par étapes, en transformant le problème de la convergence d'une suite en celui de la convergence d'une série, pour conclure en appliquant le théorème de Beppo Levi.

Soit $(f_n)_{n\in\mathbf{N}}$ une suite de Cauchy d'éléments de $\mathcal{L}^1$. Pour tout $k \geqslant 0$, il existe un entier n_k tel que

$$\forall n, m \geqslant n_k,\ \|f_n - f_m\|_{\mathrm{L}^1} \leqslant \frac{1}{2^k}.$$

On peut, sans perte de généralité et quitte à prendre $\max\{n_0, \ldots, n_k\}$, supposer que la suite n_k est strictement croissante. Posons maintenant $g_0 = f_{n_0}$ et $g_k = f_{n_k} - f_{n_{k-1}}$ pour $k \geqslant 1$. On a alors

$$\|g_k\|_{\mathrm{L}^1} \leqslant \frac{1}{2^k}$$

et donc la série $\sum \|g_k\|_{\mathrm{L}^1}$ est convergente. D'après le théorème de Beppo Levi, la série $\sum g_k$ converge presque partout vers une fonction $g \in \mathcal{L}^1(\mathbf{R}^d)$. Mais on a

$$\sum_{i=0}^{k} g_i = f_{n_k}$$

et donc (f_{n_k}) converge vers g presque partout quand k tend vers l'infini. De plus,

$$\|g - f_{n_k}\|_{\mathrm{L}^1} \leqslant \sum_{i=k+1}^{+\infty} \|f_{n_i} - f_{n_{i-1}}\|_{\mathrm{L}^1} \leqslant \frac{1}{2^{k-1}}.$$

Comme par ailleurs pour tout $m \geqslant n_k$ on a

$$\|f_m - f_{n_k}\|_{\mathrm{L}^1} \leqslant \frac{1}{2^k},$$

on en déduit que pour tout $m \geqslant n_k$ on a

$$\|f_m - g\|_{\mathrm{L}^1} \leqslant \|f_m - f_{n_k}\|_{\mathrm{L}^1} + \|f_{n_k} - g\|_{\mathrm{L}^1} \leqslant \frac{1}{2^{k-2}}$$

et donc (f_n) converge en moyenne vers g, ce qui achève la démonstration du théorème de Fischer-Riesz. □

Dans le courant de la preuve ci-dessus, nous avons en fait montré deux résultats d'intérêt indépendant :

Proposition 11.13 — *Un espace* V *muni d'une semi-norme* $\|\cdot\|$ *est complet si et seulement si toute série absolument convergente est convergente.*

Proposition 11.14 — *Si une suite* (f_n) *d'éléments de* $\mathcal{L}^1(\mathbf{R}^d)$ *converge en moyenne vers* $f \in \mathcal{L}^1(\mathbf{R}^d)$*, alors il existe une sous-suite* (f_{n_k}) *qui converge presque partout vers* f.

Insistons ici sur un fait important : en général, la suite (f_n) elle même ne converge pas presque partout vers f. Donnons un exemple de ce phénomène.

L'exemple de la bosse mouvante

Énumérons ainsi les intervalles dyadiques de $[0,1]$:

$$\mathrm{I}_1 = [0,1],\ \ \mathrm{I}_2 = [0,1/2],\ \ \mathrm{I}_3 = [1/2,1],\ \ \mathrm{I}_4 = [0,1/4],$$
$$\mathrm{I}_5 = [1/4,1/2],\ \ \mathrm{I}_6 = [1/2,3/4],\ \ \mathrm{I}_7 = [3/4,1],\ \ \mathrm{I}_8 = [0,1/8]\ldots$$

On peut en fait trouver une expression générale pour le n^{e} intervalle de cette suite, en posant

$$\mathrm{I}_n = \left[\frac{n-2^k}{2^k}, \frac{n-2^k+1}{2^k}\right], \quad \text{où } k = [\mathrm{Log}_2\, n].$$

On note alors f_n la fonction indicatrice de l'intervalle I_n. Elle vérifie

$$\|f_n\|_{\mathrm{L}^1} \leqslant \frac{1}{2^k},$$

et donc (f_n) converge en moyenne vers la fonction nulle. Mais la « bosse » que constitue le graphe de f_n se promène de droite à gauche dans l'intervalle $[0,1]$, indéfiniment. Ainsi pour tout $x \in [0,1]$, il existe une sous-suite (f_{n_i}) telle que $f_{n_i}(x) = 1$ pour tout i, et une autre sous-suite (f_{m_i}) telle que $f_{m_i}(x) = 0$. Ainsi, pour tout $x \in [0,1]$, la suite $(f_n(x))$ a deux valeurs d'adhérences, 0 et 1, et elle ne peut donc converger. Par contre, il est facile de vérifier que la sous-suite (f_{2^n}) converge vers 0 pour tout $x \neq 0$.

Espace L^p

On peut considérer les autres espaces de Lebesgue : si $p \geqslant 1$ est un nombre réel et Ω est un ouvert de $\mathbf{R}^d$, notons $\mathcal{L}^p(\Omega)$ l'espace des fonctions mesurables telles que $|f|^p$ soit intégrable sur Ω. On peut munir $\mathcal{L}^p(\Omega)$ de la *semi-norme* L^p définie par

$$\|f\|_{\mathrm{L}^p} = \left(\int_\Omega |f(\mathbf{x})| \,\mathrm{d}\mathbf{x}\right)^{1/p},$$

et c'est bien une semi-norme, dont l'inégalité triangulaire n'est autre que l'inégalité de Minkowski (*voir* ci-dessous). On peut, comme pour l'espace L^1, faire un quotient par le sous-espace des fonctions nulles presque partout pour obtenir l'espace de Lebesgue $\mathrm{L}^p(\Omega)$ et démontrer le théorème suivant :

Théorème 11.15 — *Pour tout réel $p \geqslant 1$, l'espace normé $(\mathrm{L}^p(\Omega), \|\cdot\|_{\mathrm{L}^p})$ est un espace de Banach.*

Inégalités de Cauchy-Schwarz, Hölder et Minkovski

Nous n'avons jusqu'ici démontré ces inégalités que pour les fonctions continues à support compact (en une variable, mais la démonstration est exactement la même en plusieurs). Cela fonctionne bien car le produit et la somme de deux fonctions continues à supports compacts est encore une fonction continue à support compact. Le fait que les fonctions mesurables soient également stables par ces opérations et le théorème 11.1 vont nous permettre d'étendre largement le domaine de validité de ces inégalités.

Théorème 11.16 — *Soient Ω un ouvert de $\mathbf{R}^d$ et f, g deux fonctions mesurables définies sur Ω.*

1. Inégalité de Cauchy-Schwarz : *si f et g sont de carré intégrable sur Ω, alors fg est Lebesgue-intégrable sur Ω et*

$$\int_\Omega |f(\mathbf{x})g(\mathbf{x})| \,\mathrm{d}\mathbf{x} \leqslant \left(\int_\Omega f(\mathbf{x})^2 \,\mathrm{d}\mathbf{x}\right)^{1/2} \left(\int_\Omega g(\mathbf{x})^2 \,\mathrm{d}\mathbf{x}\right)^{1/2}.$$

2. Inégalité de Hölder : *soit $p > 1$, et $q > 1$ tels que $1/p + 1/q = 1$. Si $f \in \mathrm{L}^p(\Omega)$ et $g \in \mathrm{L}^q(\Omega)$, alors $fg \in \mathrm{L}^1(\Omega)$ et*

$$\int_\Omega |f(\mathbf{x})g(\mathbf{x})| \,\mathrm{d}\mathbf{x} \leqslant \left(\int_\Omega f(\mathbf{x})^p \,\mathrm{d}\mathbf{x}\right)^{1/p} \left(\int_\Omega g(\mathbf{x})^q \,\mathrm{d}\mathbf{x}\right)^{1/q}.$$

3. Inégalité de Minkowski : *soit $p \geqslant 1$. Si $f, g \in \mathrm{L}^p(\Omega)$ alors $f + g \in \mathrm{L}^p(\Omega)$ et*

$$\left(\int_\Omega |f(\mathbf{x}) + g(\mathbf{x})|^p \,\mathrm{d}\mathbf{x}\right)^{1/p} \leqslant \left(\int_\Omega |f(\mathbf{x})|^p \,\mathrm{d}\mathbf{x}\right)^{1/p} + \left(\int_\Omega |g(\mathbf{x})|^p \,\mathrm{d}\mathbf{x}\right)^{1/p}.$$

Démonstration — Nous ne démontrerons que l'inégalité de Hölder dont les inégalités de Cauchy-Schwarz et Minkowski sont des conséquences simples. Soient $f \in \mathrm{L}^p(\Omega)$ et $g \in \mathrm{L}^q(\Omega)$ avec $p^{-1} + q^{-1} = 1$. Les fonctions f et g sont mesurables et donc fg l'est aussi.

Si a, b sont deux nombres réels et si $t > 0$, on a

$$|ab| = \exp\left(p^{-1}\operatorname{Log}(|ta|^p) + q^{-1}\operatorname{Log}(|tb|^q)\right)$$

et la convexité de la fonction exponentielle entraîne que

$$|ab| \leqslant \frac{t^p}{p}|a|^p + \frac{t^q}{q}|b|^q.$$

Soit $\mathbf{x} \in \Omega$. On a alors

$$|f(\mathbf{x})g(\mathbf{x})| \leqslant \frac{t^p}{p}|f(\mathbf{x})|^p + \frac{t^q}{q}|g(\mathbf{x})|^q$$

et donc fg est Lebesgue-intégrable sur Ω d'après le théorème 11.1. On en déduit immédiatement que

$$\int_\Omega |f(\mathbf{x})g(\mathbf{x})|\,\mathrm{d}\mathbf{x} \leqslant \frac{t^p}{p}\|f\|_{\mathrm{L}^p}^p + \frac{t^q}{q}\|g\|_{\mathrm{L}^q}^q.$$

L'avantage est que l'on peut ajuster le paramètre pour essayer d'obtenir l'inégalité optimale. Il est facile de voir que celle-ci a lieu pour $t = \|f\|_{\mathrm{L}^p}^{-1/q}\|g\|_{\mathrm{L}^q}^{-1/p}$ et que de $p^{-1} + q^{-1} = 1$ il découle

$$\|fg\|_{\mathrm{L}^1} \leqslant \|f\|_{\mathrm{L}^p}\|g\|_{\mathrm{L}^q},$$

ce qui n'est autre que l'inégalité de Hölder. □

11.6. Théorèmes de densité

Nous allons ici démontrer deux théorèmes, l'un portant sur la densité des fonctions continues et l'autre sur celle des fonctions en escalier. Munis de ces deux théorèmes, nous pourrons étendre des résultats démontrés dans le cas des fonctions continues à un cadre bien plus général, et ce (presque) sans peine.

Théorème 11.17 — *Soient Ω un ouvert de $\mathbf{R}^d$ et f un élément de $\mathcal{L}^1(\Omega)$. Il existe une suite $(g_n)_{n\in\mathbf{N}}$ de fonctions continues et intégrables sur Ω qui converge presque partout et en moyenne vers f.*

Dit autrement, l'espace $\mathcal{C}^0(\Omega) \cap \mathcal{L}^1(\Omega)$ est dense dans $\mathcal{L}^1(\Omega)$ pour la topologie définie par la norme L^1.

Démonstration — Fixons $\varepsilon > 0$ et commençons par démontrer qu'il existe une fonction continue et intégrable g telle que $\|f-g\|_{\mathrm{L}^1} \leqslant \varepsilon$. Pour tout entier $n \geqslant 0$, notons $f_n = \text{méd}\{-n, f, n\}$ la troncature de f par $-n, n$. Bien sûr f_n est Lebesgue-intégrable et, de plus, (f_n) converge simplement vers la fonction f quand n tend vers l'infini. Comme en outre $|f_n(x)| \leqslant |f(x)|$ pour tout x, le théorème de convergence dominée implique que (f_n) converge en moyenne vers f. Ainsi si l'on prend n suffisamment grand, on aura $\|f - f_n\| \leqslant \varepsilon/2$. Considérons maintenant la fonction f_n ; comme elle est mesurable, il existe une suite (h_k) de fonctions continues sur Ω qui converge presque partout vers f_n. Quitte à changer h_k en $\text{méd}\{-n, h_k, n\}$ [1], on peut supposer que $|h_k(x)| \leqslant n$ pour tout $k \geqslant 0$. Si Ω est borné, alors la fonction constante égale à n est Lebesgue-intégrable sur Ω et l'on peut appliquer le théorème de convergence dominée à la suite (h_k) et donc pour k suffisamment grand on a $\|h_k - g_n\| \leqslant \varepsilon/2$. On en déduit que $\|f - h_k\| \leqslant \|f - g_n\| + \|g_n - h_k\| \leqslant \varepsilon$, ce qui démontre l'assertion. Si Ω n'est pas borné, il faut une étape supplémentaire consistant à prendre des unions finies de pavés inclus dans Ω qui vont couvrir une partie de plus en plus important de Ω, puis appliquer le théorème de Hake. Nous laissons au lecteur scrupuleux la mise en place de ce dernier point.

Nous pouvons maintenant achever la démonstration du théorème : pour tout entier $n \geqslant 1$ il existe une fonction f_n continue et intégrable sur Ω telle que $\|f - f_n\|_{\mathrm{L}^1} \leqslant 1/n$, et donc f est la limite en moyenne de la suite (f_n) ; on sait alors qu'il existe une sous-suite de (f_n) qui, en plus, converge presque partout vers f. □

Il est parfois plus facile de travailler avec des fonctions en escalier qu'avec des fonctions continues. Afin de simplifier l'exposition, nous ne démontrerons la densité des fonctions en escalier dans $\mathcal{L}^1$ qu'en dimension 1 et sur un intervalle compact. Un théorème analogue est vrai en toute dimension et sur des domaines quelconques, pour toute définition raisonnable de fonction en escalier, mais nous laisserons cela de côté. L'étape délicate est de montrer une caractérisation des fonctions mesurables en termes de fonctions en escalier :

Proposition 11.18 — *Une fonction $f : [a, b] \longrightarrow \mathbf{R}$ est mesurable si et seulement si elle est limite simple presque partout d'une suite de fonctions en escalier.*

Démonstration — Commençons par démontrer que la condition d'être limite d'une suite de fonctions en escalier est nécessaire. Soit donc f une fonction mesurable et (f_n) une suite de fonctions continues sur $[a, b]$ qui converge presque partout vers f. Notons $\mathrm{Z} \subset [a, b]$ l'ensemble de mesure nulle des x tels que $(f_n(x))$ ne converge pas vers $f(x)$. Une fonction continue est limite uniforme d'une suite de fonctions en escalier : on dit que c'est une fonction réglée. Pour tout entier $n \geqslant 1$ il existe donc une fonction en escalier h_n telle que

$$\forall x \in [a, b], \quad |f_n(x) - h_n(x)| \leqslant \frac{1}{n}.$$

1. Si a, b, c sont trois nombres réels, $\text{méd}\{a, b, c\}$ est celui des trois nombres situé entre les deux autres.

Pour tout $x \in [a,b]$ on a alors

$$|f(x) - h_n(x)| \leqslant |f(x) - f_n(x)| + 1/n$$

et donc pour tout $x \notin \mathrm{Z}$ on a $\lim_{n\to+\infty} h_n(x) = f(x)$, ce qu'il fallait démontrer.

Passons maintenant à la suffisance, et soit donc (h_n) une suite de fonctions en escalier qui converge presque partout vers f ; comme précédemment, notons Z l'ensemble de mesure nulle où la convergence n'a pas lieu. Nous pouvons sans perte de généralité supposer que h_n est continue aux points a et b, et alors pour tout n la fonction h_n est discontinue en un nombre fini de points que l'on peut inclure dans une union disjointe finie d'intervalles ouverts $\mathrm{J}_1^n, \ldots, \mathrm{J}_{l(n)}^n$, dont la somme des longueurs est inférieure à 2^{-n}. Soit f_n la fonction continue sur $[a,b]$ qui soit égale à h_n hors des intervalles J_l^n et qui, sur un tel intervalle, soit affine (*voir* la figure 11.1).

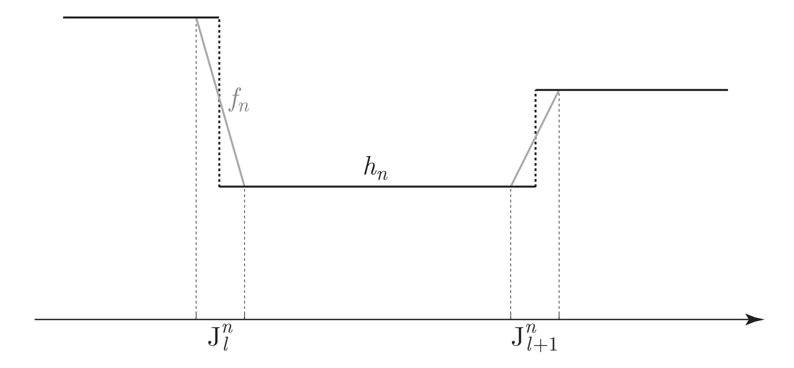


Figure 11.1 – Approximation d'une fonction en escalier par une fonction affine par morceaux

On pose maintenant

$$\mathrm{E}_n = \mathrm{J}_1^n \cup \cdots \cup \mathrm{J}_{l(n)}^n.$$

Si $x \in [a,b] - \mathrm{Z}$ n'est que dans un nombre fini de E_n, alors $f_n(x) = h_n(x)$ pour tout x assez grand et donc $\lim_{n\to+\infty} f_n(x) = f(x)$. Mais comme la série $\sum |\mathrm{E}_n|$ converge, le lemme de Borel-Cantelli nous assure que l'ensemble $\mathrm{E}^\star$ des x appartenant à une infinité de E_n est de mesure nulle, et donc pour tout $x \in [a,b] - (\mathrm{E}^\star \cup \mathrm{Z})$, on a $\lim_{n\to+\infty} f_n(x) = f(x)$, ce qui achève la démonstration du théorème. $\square$

On peut en fait démontrer l'analogue de ce résultat en toute dimension et sur tout domaine. Une fonction en escalier sur un ouvert Ω de $\mathbf{R}^d$ est une combinaison linéaire (en particulier une somme finie) de fonctions indicatrices de pavés compacts inclus dans Ω. C'est la construction de l'approximation d'une fonction en escalier par une fonction continue qui est un peu plus technique. Mais une fois acquis ce résultat, on peut, exactement comme on l'a fait pour les fonctions continues, démontrer le théorème de densité suivant.

Théorème 11.19 — *Soient Ω un ouvert de $\mathbf{R}^d$ et f un élément de $\mathcal{L}^1(\mathbf{R}^d)$. Alors il existe une suite $(h_n)_{n\in\mathbf{N}}$ de fonctions en escalier intégrables sur Ω qui converge en moyenne et presque partout vers f.*

Il n'est pas difficile, une fois démontrés les résultats ci-dessus, de démontrer que les ensembles de fonctions plus régulières que simplement continues sont aussi denses dans les espaces L^p, et qu'on peut les supposer à support compact (rappelons que le support d'une fonction est l'adhérence de l'ensemble des points où f est non nulle).

Théorème 11.20 — *Soient Ω un ouvert de $\mathbf{R}^d$, $p \in [1, +\infty[$ et $k \in \mathbf{N} \cup \{\infty\}$. Toute fonction $f \in \mathrm{L}^p(\Omega)$ est limite en norme L^p et presque partout d'une suite de fonctions de classe $\mathcal{C}^k$ à supports compacts dans Ω.*

Une application classique des théorèmes de densité

Finissons cette section par une application des théorèmes de densité qui illustre parfaitement l'usage que l'on fait de ces résultats. Nous allons démontrer la forme générale du **théorème de changement de variable**.

Soient Ω un ouvert de $\mathbf{R}^d$, $\varphi : \Omega \longrightarrow \varphi(\Omega)$ un difféomorphisme de classe $\mathcal{C}^1$ et f une fonction Lebesgue-intégrable sur $\varphi(\Omega)$. Nous voulons démontrer la formule du jacobien, à savoir

$$\int_{\varphi(\Omega)} f(\mathbf{y})\,\mathrm{d}\mathbf{y} = \int_\Omega f(\varphi(\mathbf{x}))\,|\operatorname{Jac}\varphi(\mathbf{x})|\,\mathrm{d}\mathbf{x}.$$

La fonction f n'est définie que sur $\varphi(\Omega)$ privé d'un ensemble de mesure nulle Z. Ainsi $f \circ \varphi$ n'est définie que sur $\Omega - \varphi^{-1}(\mathrm{Z})$ et il nous faut donc vérifier que $\varphi^{-1}(\mathrm{Z})$ est de mesure nulle. C'est le contenu du lemme suivant, dont l'énoncé est intéressant en soi :

Lemme 11.21 — *Soient U, V deux ouverts de $\mathbf{R}^d$ et ψ un difféomorphisme de classe $\mathcal{C}^1$ de U sur V. Si $\mathrm{Z} \subset \mathrm{U}$ est de mesure nulle, alors $\psi(\mathrm{Z})$ l'est aussi.*

Démonstration — Par densité de $\mathbf{Q}^d$ dans $\mathbf{R}^d$, pour tout $r > 0$ l'ouvert U est réunion dénombrable de boules de rayon au plus r. Comme une union dénombrable d'ensembles de mesure nulle est de mesure nulle, il nous suffit de démontrer le théorème dans le cas suivant : U est la boule ouverte $\mathrm{B}_<(0, r)$ de rayon $r > 0$, Z est inclus dans la boule fermée $\mathrm{B}_\leqslant(0, r')$ de rayon $0 < r' < r$ et ψ est défini sur U. Fixons $\varepsilon > 0$. Comme Z est de mesure nulle, il existe une famille dénombrable de boules $(\mathrm{B}_n)_{n\in\mathbf{N}}$ telle que

$$\mathrm{Z} \subset \bigcup_{n\in\mathbf{N}} \mathrm{B}_n \quad \text{et} \quad \sum_{n=0}^{\infty} |\mathrm{B}_n| \leqslant \varepsilon.$$

Comme $\mathrm{Z} \subset \mathrm{B}_\leqslant(0, r')$ et que $r' < r$, on peut supposer que toutes les boules B_n sont incluses dans une même boule $\mathrm{B}'' = \mathrm{B}_\leqslant(o, r'')$ avec $r'' < r$. Sur cette boule, qui est compacte, l'inégalité des accroissements finis nous assure qu'il existe une constante

$\lambda > 0$ telle que, pour tout $\mathbf{x}, \mathbf{y} \in \mathrm{B}''$ on ait $\|\psi(\mathbf{x}) - \psi(\mathbf{y})\| \leqslant \|\mathbf{x} - \mathbf{y}\|$. Ainsi, pour tout $n \geqslant 0$, l'image de la boule B_n est incluse dans une boule B'_n dont le rayon est égal à celui de B_n multiplié par λ. Nous en déduisons que

$$\psi(\mathrm{Z}) \subset \bigcup_{n \in \mathbf{N}} \mathrm{B}'_n \quad \text{et} \quad \sum_{n=0}^{\infty} |\mathrm{B}'_n| \leqslant \lambda^d \varepsilon,$$

ce qui démontre le lemme : $\psi(\mathrm{Z})$ est de mesure nulle. □

Continuons notre preuve du cas général du théorème de changement de variable. D'après le théorème de densité, il existe une suite $(f_n)_{n \in \mathbf{N}}$ de fonctions continues à support compact dans Ω telles que, pour tout $n \geqslant 0$, on ait $\|f - f_n\|_{\mathrm{L}^1} \leqslant 2^{-n}$. Appliquons maintenant une astuce qu'il est utile de connaître et que nous avons déjà utilisée : définissons la suite de fonctions (g_k) qui soit en quelque sorte la primitive discrète de la suite (f_k). On pose $g_0 = f_0$ puis, pour tout $k \geqslant 1$, $g_k = f_k - f_{k-1}$. L'avantage est qu'alors f s'écrit comme la somme d'une série convergente dans L^1 :

$$f = \sum_{k=0}^{\infty} g_k.$$

Décomposons maintenant chaque fonction f, g_k en ses parties positives et négatives, de sorte que

$$f^+ = \sum_{k=0}^{\infty} g_k^+.$$

La fonction f^+ apparaît donc comme une limite *croissante* de fonctions continues, de même que la fonction f^- apparaît comme la limite décroissante d'une suite de fonctions continues. Nous pouvons donc, par linéarité de l'intégrale, nous contenter de démontrer la formule du Jacobien *pour une fonction Lebesgue-intégrable f qui est en norme* L^1 *la limite croissante d'une suite de fonctions continues à support compact* f_k^+. D'après le théorème de Beppo Levi, la fonction f est alors aussi limite presque partout de la suite de fonctions f_k^+ et l'on déduit de la formule du changement de variable pour les fonctions continues que

$$\lim_{k \to +\infty} \int_{\varphi(\Omega)} f_k^+ \circ \varphi |\operatorname{Jac} \varphi(\mathbf{x})| \, \mathrm{d}\mathbf{x} = \lim_{k \to +\infty} \int_{\Omega} f_k^+(\mathbf{y}) \, \mathrm{d}\mathbf{y} = \int_{\Omega} f(\mathbf{y}) \, \mathrm{d}\mathbf{y},$$

ce qui achève la démonstration du théorème de changement de variable pour les fonctions Lebesgue-intégrables.

11.7. La théorie L^2 des séries de Fourier

L'espace L^2, comme les espaces L^p, est un très bel espace vectoriel normé complet. Mais il y a mieux : on peut le munir d'un produit scalaire et donc y parler d'orthogonalité, en particulier de base orthonormée. Nous revenons dans cette section sur la théorie

des séries de Fourier pour lesquelles nous allons exploiter ce que nous connaissons maintenant des espaces de Lebesgue.

11.7.1. Géométrie de l'espace hilbertien L^2

Nous nous intéresserons ici à l'espace des (classes de) fonctions d'une variable réelle, à valeurs complexes, qui sont 2π-périodiques et dont le carré est Lebesgue-intégrable sur $[0, 2\pi]$. Nous noterons cet espace $\mathrm{L}^2(\mathbf{T})$, où $\mathbf{T}$ désigne le cercle unité de $\mathbf{C}$. En effet, il existe une bijection naturelle entre l'espace des fonctions 2π-périodique et celui des fonctions définies sur le cercle. Nous penserons cependant à ces fonctions comme étant des fonctions de la variable réelle. Une fonction f est élément de $\mathrm{L}^2(\mathbf{T})$ si et seulement si f est mesurable et si $|f|^2$ est intégrable sur $[0, 2\pi]$. Comme mentionné dans la section sur les théorèmes de complétude, nous définissons la norme L^2 de f par

$$\|f\|_2 = \left(\frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2 \,\mathrm{d}x\right)^{1/2} = \left(\frac{1}{2\pi}\int_0^{2\pi} f(x)\overline{f(x)} \,\mathrm{d}x\right)^{1/2}.$$

C'est bien une norme sur l'espace $\mathrm{L}^2(\mathbf{T})$ des classes de fonctions et pour cette norme l'espace est complet. Mais le second membre de l'égalité ci-dessus nous incite à définir, en plus de la norme, un *produit scalaire* sur l'espace $\mathrm{L}^2(\mathbf{T})$. Si $f, g \in \mathrm{L}^2(\mathbf{T})$, alors f et g sont mesurables et donc fg l'est aussi. De l'inégalité élémentaire

$$2|f\overline{g}| = 2|f||g| \leqslant |f|^2 + |g|^2$$

on déduit que $f\overline{g}$ est Lebesgue-intégrable.

Définition 11.7 — *Si f, g sont deux éléments de $\mathrm{L}^2(\mathbf{T})$, leur* produit scalaire L^2 *est défini par*

$$(f, g) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)} \,\mathrm{d}t.$$

L'intégrale peut en fait être prise sur tout intervalle de longueur 2π.

Si nous avons choisi le nom de produit scalaire, c'est que cela en est bien un. Il est facile de démontrer la proposition suivante, dont nous laissons la preuve au lecteur.

Proposition 11.22 — *Le produit scalaire L^2 défini ci-dessus munit $\mathrm{L}^2(\mathbf{T})$ d'une forme hermitienne au sens où les propriétés suivantes sont satisfaites pour toutes $f, g, h \in \mathrm{L}^2(\mathbf{T})$ et tout $\lambda \in \mathbf{C}$:*

1. $(f, g) = \overline{(g, f)}$;
2. $(f + g, h) = (f, h) + (g, h)$;
3. $(\lambda f, g) = \lambda(f, g)$;
4. $(f, f) \geqslant 0$;
5. $(f, f) = 0$ *si et seulement si* $f = 0$.

La dernière assertion de la proposition n'est vraie que parce que nous considérons les classes de fonctions modulo égalité presque partout. Si f est une « vraie » fonction, cette propriété s'énoncerait ainsi :

$$(f, f) = 0 \text{ si et seulement si } f(x) = 0 \text{ p.p.}$$

Nous ferons souvent usage des deux inégalités de Schwarz et Minkowski, que nous rappelons ici :

$$|(f, g)| \leqslant \|f\|_2 \|g\|_2 \text{ et } \|f + g\|_2 \leqslant \|f\|_2 + \|g\|_2.$$

La norme de l'espace $\mathrm{L}^2(\mathbf{T})$ est donc définie par une forme hermitienne. L'espace normé ainsi obtenu étant en plus complet, on dit que $\mathrm{L}^2(\mathbf{T})$ est un *espace de Hilbert.* Ces espaces jouent un rôle éminent dans tous les domaines des mathématiques où l'analyse intervient, de la théorie de la représentation des groupes à la géométrie différentielle en passant par l'analyse appliquée, le traitement du signal ou la mécanique quantique. Et s'ils jouent un rôle aussi important, c'est parce qu'ils sont le lieu le plus agréable où faire de l'analyse de Fourier !

Pour tout $n \in \mathbf{Z}$, la fonction $\mathrm{e}_n : x \mapsto \mathrm{e}^{inx}$ est dans $\mathrm{L}^2(\mathbf{T})$. Un calcul élémentaire montre que $(\mathrm{e}_n, \mathrm{e}_n) = 1$ et que si $n \neq m$ alors $(\mathrm{e}_n, \mathrm{e}_m) = 0$. La famille $\{\mathrm{e}_n, n \in \mathbf{Z}\}$ forme donc un *système orthonormé* de l'espace de Hilbert $\mathrm{L}^2(\mathbf{T})$. C'est en particulier une famille libre d'éléments de la sphère unité de $\mathrm{L}^2(\mathbf{T})$. Un calcul tout aussi élémentaire démontre que si $n \neq m$ alors $\|\mathrm{e}_n - \mathrm{e}_m\|_2 = \sqrt{2}$ de sorte qu'aucune sous-suite de la suite (e_n) ne converge dans $\mathrm{L}^2(\mathbf{T})$. Mais il y a mieux encore ; si f est dans l'espace vectoriel engendré par les e_n, nous pouvons le décomposer en combinaison linéaire de la forme $f = \sum \lambda_n \mathrm{e}_n$, la somme étant ici finie, avec $\lambda_n \in \mathbf{C}$. Dans un espace muni d'un produit scalaire, les coordonnées d'un vecteur dans une base orthonormée s'expriment par un produit scalaire, d'où l'on déduit

$$c_n = (f, \mathrm{e}_n) = \frac{1}{2\pi} \int_0^{2\pi} f(x) \mathrm{e}^{-inx} \, \mathrm{d}x,$$

de sorte que les coordonnés de f dans la base $\{\mathrm{e}_n, n \in \mathbf{Z}\}$ ne sont autres que *les coefficients de Fourier de f.*

Le résultat central de cette section est que, en un sens à préciser, la famille $\{\mathrm{e}_n, n \in \mathbf{Z}\}$ forme une base de tout l'espace $\mathrm{L}^2(\mathbf{T})$.

Définition 11.8 — *La famille de vecteurs $\{f_n, n \in \mathbf{N}\}$ de l'espace de Hilbert $\mathrm{L}^2(\mathbf{T})$ est une* base hilbertienne *de $\mathrm{L}^2(\mathbf{T})$ si la famille est orthonormée et si, pour tout $f \in \mathrm{L}^2(\mathbf{T})$, il existe une suite (c_n) de nombres complexes telle que les deux conditions suivantes soient satisfaites :*

1. La suite (c_n) est de carré sommable :

$$\sum_{n \in \mathbf{N}} |c_n|^2 < +\infty\,;$$

2. La suite de (classes de) fonctions

$$\sum_{k=0}^{n} c_n f_n$$

converge en norme L^2 *vers la (classe de) fonction* f.

Lorsque la seconde condition est vérifiée, on note simplement

$$f = \sum_{k=0}^{\infty} c_n f_n.$$

On peut démontrer que, la famille (f_n) étant orthonormée, elle est libre et que, de plus, les coefficients de f dans la base hilbertienne (f_n) s'expriment au moyen d'un produit scalaire : $c_n = (f, f_n)$. Enfin on généralise sans peine la définition ci-dessus au cas où les indices n varient dans $\mathbf{Z}$ plutôt que dans $\mathbf{N}$.

Tout espace de Hilbert n'admet pas nécessairement de base hilbertienne. Il est assez facile de démontrer que $\mathrm{L}^2(\mathbf{T})$ en admet une, car il est *séparable.*

Proposition 11.23 — *Il existe une famille dénombrable* $\{g_n, n \in \mathbf{N}\}$ *d'éléments de* $\mathrm{L}^2(\mathbf{T})$ *tel que l'espace vectoriel qu'elle engendre est* dense *dans* $\mathrm{L}^2(\mathbf{T})$. *On dit que* $\mathrm{L}^2(\mathbf{T})$ *est un espace de Hilbert séparable.*

La démonstration est simple et nous en laissons les détails au lecteur : on peut utiliser la densité des fonctions en escalier dans $\mathrm{L}^2(\mathbf{T})$ et s'apercevoir que l'on peut ne considérer que les fonctions en escalier dont les points de discontinuité et les valeurs sont des nombres rationnels. En appliquant le procédé d'orthogonalisation de Gram-Schmidt à la famille $\{g_n\}$ on obtient le corollaire suivant :

Corollaire 11.24 — *L'espace de Hilbert* $\mathrm{L}^2(\mathbf{T})$ *admet une base hilbertienne.*

Démontrons maintenant une inégalité de nature géométrique qui nous sera utile par la suite.

Proposition 11.25 — *Soit* $\{f_n, n \in \mathbf{N}\}$ *une famille orthonormée d'éléments de* $\mathrm{L}^2(\mathbf{T})$. *Soient* $f \in \mathrm{L}^2(\mathbf{T})$ *et posons* $c_k = (f, f_k)$ *pour tout* $k \in \mathbf{N}$. *Pour tout entier* $n \geqslant 0$ *et toute famille* $\lambda_0, \ldots, \lambda_n$ *de nombres complexes, on a*

$$\left\| f - \sum_{k=0}^{n} c_k f_k \right\|_2 \leqslant \left\| f - \sum_{k=0}^{n} \lambda_k f_k \right\|_2$$

et il y a égalité si et seulement si $\lambda_0 = c_0, \ldots, \lambda_n = c_n$. *Dit autrement,* $c_0 f_0 + \cdots + c_n f_n$ *réalise l'unique minimum de la distance* L^2 *entre* f *et les éléments du sous-espace engendré par* $\{f_0, \ldots, f_n\}$.

Démonstration — Nous allons, comme souvent, utiliser l'identité $a = a + b - b$ et la bilinéarité du produit scalaire :

$$\begin{aligned}\left\| f - \sum_{k=0}^{n} \lambda_k f_k \right\|_2^2 &= \left\| f - \sum_{k=0}^{n} c_k f_k + \sum_{k=0}^{n} (c_k - \lambda_k) f_k \right\|_2^2 \\ &= \left\| f - \sum_{k=0}^{n} c_k f_k \right\|_2^2 \\ &\qquad + 2\mathrm{Re}\left(f - \sum_{k=0}^{n} c_k f_k, \sum_{k=0}^{n} (c_k - \lambda_k) f_k \right) \\ &\qquad + \left\| \sum_{k=0}^{n} (c_k - \lambda_k) f_k \right\|_2^2 .\end{aligned}$$

Comme $c_k = (f, f_k)$, le produit scalaire sous la partie réelle dans l'égalité ci-dessus est nul car la famille des f_k est orthogonale. Il en ressort que

$$\left\| f - \sum_{k=0}^{n} \lambda_k f_k \right\|_2^2 = \left\| f - \sum_{k=0}^{n} c_k f_k \right\|_2^2 + \left\| \sum_{k=0}^{n} (c_k - \lambda_k) f_k \right\|_2^2 ,$$

ce qui achève la démonstration de la proposition. □

Il est alors aisé de démontrer le corollaire suivant, connu sous le nom *d'inégalité de Bessel :*

Corollaire 11.26 — *Avec les notations de la proposition 11.25, on a l'inégalité*

$$\sum_{k=0}^{\infty} |c_k|^2 \leqslant \|f\|_2^2 .$$

Si de plus $\{f_n, n \in \mathbf{N}\}$ est une base hilbertienne, nous obtenons un résultat plus précis, connu sous le nom *d'identité de Plancherel* :

Corollaire 11.27 — *Soit $\{f_n, n \in \mathbf{N}\}$ une base hilbertienne de l'espace $\mathrm{L}^2(\mathbf{T})$. Si $f \in \mathrm{L}^2(\mathbf{T})$ et si, pour tout $k \in \mathbf{N}$ on pose $c_k = (f, f_k)$, alors*

$$\sum_{k=0}^{+\infty} |c_k|^2 = \|f\|_2^2 .$$

Nous concluons cette section finalement très générale sur la géométrie des espaces de Hilbert séparables par le théorème connu sous le nom de théorème de Fischer-Riesz. Nous devons pour cela introduire l'espace $\ell^2(\mathbf{N})$ des suites de nombres complexes de carré sommable.

Définition 11.9 — *On note $\ell^2(\mathbf{N})$ l'ensemble de toutes les suites $(u_n)_{n\in\mathbf{N}}$ de nombres complexes qui satisfont à*

$$\sum_{n=0}^{\infty} |u_n|^2 < \infty.$$

Si $\mathrm{U} = (u_n)$ et $\mathrm{V} = (v_n)$ sont deux éléments de $\ell^2(\mathbf{N})$, la série de terme général $u_n\overline{v_n}$ est convergente et si l'on pose

$$(\mathrm{U}, \mathrm{V}) = \sum_{n=0}^{\infty} u_n\overline{v_n},$$

on munit $\ell^2(\mathbf{N})$ d'une structure d'espace de Hilbert séparable dont une base hilbertienne est donnée par la famille $\{\Delta_k, k \in \mathbf{N}\}$ où $\Delta_k = (\delta_{k,n})_{n\in\mathbf{N}}$.

Le théorème de Fischer-Riesz affirme simplement que *tout espace de Hilbert séparable est isométrique à l'espace $\ell^2(\mathbf{N})$.* C'est la version hilbertienne de dimension infinie d'un fait bien connu pour les espace vectoriels hermitiens de dimension finie.

Théorème 11.28 — *Soient $f \in \mathrm{L}^2(\mathbf{T})$ et $\{f_n, n \in \mathbf{N}\}$ une base hilbertienne de $\mathrm{L}^2(\mathbf{T})$. Si l'on note $c_n = (f, f_n)$ les coordonnées de f dans la base, l'application $f \mapsto \widehat{f} = (c_n)_{n\in\mathbf{N}}$ est une isomorphisme d'espaces de Hilbert entre $\mathrm{L}^2(\mathbf{T})$ et $\ell^2(\mathbf{N})$ au sens où c'est un isomorphisme linéaire qui préserve les normes.*

La démonstration de ce théorème est simple et laissée en exercice : c'est une conséquence simple de l'identité de Plancherel et de la complétude des espaces considérés.

11.7.2. Convergence uniforme de la série de Fourier d'une fonction de classe $\mathcal{C}^1$

Comme nous l'avons annoncé dans le chapitre sur les séries de Fourier, nous allons démontrer le théorème suivant :

Théorème 11.29 — *Si f est une fonction 2π-périodique de classe $\mathcal{C}^1$, la série de Fourier de f converge uniformément vers f.*

Démonstration — Notons $\mathrm{S}_n(f)$ la somme partielle de rang n de la série de Fourier de f et $c_n(f)$ le n^{e} coefficient de Fourier de f. Comme f est de classe $\mathcal{C}^1$, nous pouvons utiliser la formule d'intégration par parties pour déterminer les coefficients de Fourier de f' :

$$c_n(f') = \frac{1}{2\pi}\int_0^{2\pi} f'(x)\mathrm{e}^{-inx}\,\mathrm{d}x = \frac{in}{2\pi}\int_0^{2\pi} f(x)\mathrm{e}^{-inx}\,\mathrm{d}x = inc_n(f).$$

D'après l'inégalité triangulaire, nous obtenons, pour tous entiers m, n tels que $m \leqslant n$ et tout $x \in \mathbf{R}$:

$$|\mathrm{S}_n(f)(x) - \mathrm{S}_m(f)(x)| \leqslant \sum_{|k|>n} |c_k(f)| = \sum_{|k|>n} |k|^{-1}c_k(f').$$

Par l'inégalité de Cauchy-Schwarz,

$$|\mathrm{S}_n(f)(x) - \mathrm{S}_m(f)(x)| \leqslant \Bigg(\sum_{|k|>n} |c_k(f')|^2 \Bigg)^{1/2} \Bigg(\sum_{|k|>n} |k|^{-2} \Bigg)^{1/2}.$$

En utilisant maintenant l'inégalité de Bessel pour majorer le premier terme du produit au second membre, nous obtenons

$$|\mathrm{S}_n(f)(x) - \mathrm{S}_m(f)(x)| \leqslant \|f'\|_2 \sqrt{2} \Bigg(\sum_{k=n}^{\infty} \frac{1}{k^2} \Bigg)^{1/2}.$$

Mais la série $\sum k^{-2}$ converge et donc son reste tend vers 0. Si ε est donné, il existe donc un entier N tel que, pour tout $n \geqslant \mathrm{N}$ on ait

$$|\mathrm{S}_n(f)(x) - \mathrm{S}_m(f)(x)| < \varepsilon,$$

ce qui montre que la suite de fonctions $(\mathrm{S}_n(f))$, étant uniformément de Cauchy, doit converger uniformément. Mais nous avons démontré au chapitre IX que, f étant de classe $\mathcal{C}^1$, elle converge simplement vers la fonction f. La série de Fourier de f converge donc uniformément vers f et le théorème est démontré. □

11.7.3. Séries de Fourier des fonctions de L^2

Comme nous l'avons vu un peu plus haut, on dispose dans l'espace $\mathrm{L}^2(\mathbf{T})$ d'une famille orthonormée privilégiée $\{\mathrm{e}_n, n \in \mathbf{Z}\}$ donnée par les fonctions trigonométriques élémentaires

$$\mathrm{e}_n(x) = \mathrm{e}^{inx}.$$

Si $f \in \mathrm{L}^2(\mathbf{T})$, nous noterons dorénavant ses coefficients de Fourier complexes par

$$\widehat{f}(n) = (f, \mathrm{e}_n) = \frac{1}{2\pi} \int_0^{2\pi} f(x) \mathrm{e}^{inx}\, \mathrm{d}x$$

et $\mathrm{S}_n(f)$ la somme partielle de rang n de sa série de Fourier complexe :

$$\mathrm{S}_n(f) = \sum_{k=-n}^{n} \widehat{f}(k) \mathrm{e}^{ikx}.$$

Soit $\varepsilon > 0$; les fonctions 2π-périodiques de classe $\mathcal{C}^1$ étant denses dans $\mathrm{L}^2(\mathbf{T})$, il existe une fonction f_ε de classe $\mathcal{C}^1$ telle que $\|f - f_\varepsilon\|_2 < \varepsilon$. D'après le théorème 11.29, la série de Fourier de f_ε converge uniformément vers f_ε. Il existe donc un entier $\mathrm{N} \geqslant 0$ tel que, pour tout $n \geqslant \mathrm{N}$ on ait $\|f_\varepsilon - \mathrm{S}_n(f_\varepsilon)\|_2 < \varepsilon$.

Par linéarité, nous avons par ailleurs

$$\|\mathrm{S}_n(f) - \mathrm{S}_n(f_\varepsilon)\|_2 = \Bigg\| \sum_{k=-n}^{n} (\widehat{f}(k) - \widehat{f}_\varepsilon(k)) \mathrm{e}_n \Bigg\|_2 = \Bigg\| \sum_{k=-n}^{n} (\widehat{f - f_\varepsilon}(k)) \mathrm{e}_n \Bigg\|_2,$$

d'où l'on tire que

$$\|\mathrm{S}_n(f) - \mathrm{S}_n(f_\varepsilon)\|_2 = \left(\sum_{k=-n}^{n} |\widehat{f - f_\varepsilon}(k)|^2\right)^{1/2} \leqslant \left(\sum_{k=-\infty}^{+\infty} |\widehat{f - f_\varepsilon}(k)|^2\right)^{1/2}.$$

Le dernier membre de l'inégalité peut être majoré grâce à l'inégalité de Bessel, ce qui donne finalement :

$$\|\mathrm{S}_n(f) - \mathrm{S}_n(f_\varepsilon)\|_2 \leqslant \|f - f_\varepsilon\|_2.$$

Si maintenant nous choisissons $n \geqslant \mathrm{N}$, nous obtenons

$$\|f - \mathrm{S}_n(f)\|_2 \leqslant \|f - f_\varepsilon\|_2 + \|f_\varepsilon - \mathrm{S}_n(f_\varepsilon)\|_2 + \|\mathrm{S}_n(f_\varepsilon) - \mathrm{S}_n(f)\|_2 \leqslant 3\varepsilon$$

et nous avons démontré qu'en norme L^2, la série de Fourier de f converge vers f. Ce résultat mérite d'être résumé dans un théorème.

Théorème 11.30 — *La famille des exponentielles complexes* $\{\mathrm{e}_n(x) = \mathrm{e}^{inx}, n \in \mathbf{Z}\}$ *est une base hilbertienne de l'espace* $\mathrm{L}^2(\mathbf{T})$. *Dit autrement, pour toute fonction* $f \in \mathrm{L}^2(\mathbf{T})$, *la série de Fourier de* f *converge en norme* L^2 *vers* f. *L'application qui à une fonction* f *associe la suite* $(\widehat{f}(n))_{n\in\mathbf{Z}}$ *de ses coefficients de Fourier est un isomorphisme d'espaces de Hilbert entre* $\mathrm{L}^2(\mathbf{T})$ *et* $\ell^2(\mathbf{Z})$.

Un problème se pose alors à nous : la série de Fourier d'une fonction L^2 converge-t-elle ponctuellement vers f ? La réponse à cette question constitue l'un des grands théorèmes d'analyse du XXe siècle. Il est dû au mathématicien Lennart Carleson qui, en 1966, a démontré que la série de Fourier d'une fonction L^2 converge presque partout vers cette fonction. L'un des outils de la preuve est l'inégalité maximale de Hardy et Littlewood qui fait l'objet d'un problème dans le chapitre d'exercices à la fin de ce livre. La caractérisation des coefficients de Fourier des fonctions L^1 ou même continues est un problème ardu.

11.8. Appendice : transformations ergodiques et inégalités maximales

De nombreux problèmes de mathématiques ou de ses applications amènent à l'étude de l'itération d'une transformation f d'un espace X. Si l'on s'intéresse par exemple à l'écriture d'un nombre réel sous forme de fraction continue on tombe naturellement sur la transformation $\mathrm{G}(x) = \{x^{-1}\}$ de l'intervalle $]0, 1]$ dans lui-même (où $\{t\}$ désigne la partie fractionnaire $t - [t]$ du réel t) et l'on doit alors étudier l'orbite de x sous l'action de G, c'est-à-dire l'ensemble $x, \mathrm{G}(x), \mathrm{G} \circ \mathrm{G}(x) = \mathrm{G}^2(x), \ldots$ Si au contraire on s'intéresse à l'écriture d'un nombre réel en base 10 on rencontre immédiatement la transformation $[10](x) = \{10x\}$ de $[0, 1]$ dans lui-même, que l'on doit encore une fois itérer. D'innombrables questions se posent alors : existe-t-il des points dont l'orbite

est finie ? Des points dont l'orbite est dense ? Que dire d'un point « pris au hasard » ? La théorie de l'itération étant un sujet très vaste, nous allons nous concentrer sur un cas très particulier, à savoir celui d'une transformation d'un intervalle dans lui-même et qui a la propriété de *préserver la mesure de Lebesgue.*

Définition 11.10 — *Une transformation* T *d'un intervalle* $[a, b]$ *dans lui-même est dite* préserver la mesure de Lebesgue *si pour tout sous-ensemble mesurable* X *de* $[a.b]$ *on a* $|T^{-1}(X)| = |X|$.

La transformation $[10] : [0, 1] \longrightarrow [0, 1]$ définie ci-dessus préserve la mesure de Lebesgue. Si l'on prend un petit intervalle I alors sa préimage est constituée des 10 petits intervalles de longueur $|I|/10$. Nous verrons plus loin que la transformation G ne préserve pas la mesure de Lebesgue, mais presque. Le simple fait de préserver la mesure de Lebesgue pour T permet de répondre assez précisément à la dernière question posée ci-dessus. De manière plus générale, nous pouvons dire que nous allons en tirer des informations de nature probabiliste ou statistique sur le comportement des orbites de points « pris au hasard ». Cela signifiera en particulier que nous aurons des informations sur presque tout point. En contrepartie cela ne nous dira rien sur le comportement de l'orbite d'un point particulier qui pourrait nous intéresser.

Plaçons-nous dans l'optique suivante : l'intervalle $[a, b]$ est *l'espace des phases* d'un *système dynamique* dont l'évolution au cours du temps (discret) est donnée par la transformation T. Une fonction φ (mesurable, mais souvent plus régulière) définie sur $[0, 1]$ porte alors le nom *d'observable.* Un observateur va alors enregistrer l'évolution de l'observable φ au cours du temps, en enregistrant les valeurs successives $\varphi(x)$ (au temps 0) puis $\varphi(T(x))$ (au temps 1) puis $\varphi(T^2(x))$ (au temps 2), etc. Souvent et même avec des transformations très simples la suite de ces valeurs ne semble suivre aucun motif déterminé et en particulier elle ne converge pas. On se contente alors de considérer la moyenne temporelle des valeurs de la suite jusqu'au temps n, ce qui n'est autre que considérer les moyennes de Césaro :

$$S_n(\varphi)(x) = \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k(x)).$$

Dans le monde de la dynamique ces quantités portent le nom de *moyennes de Birkhoff,* du nom de George David Birkhoff (1884–1944), pionnier de l'étude des systèmes dynamiques. Le théorème ergodique de Birkhoff stipule que pour presque tout x (c.-à-d. pour une condition initiale prise « au hasard »), la moyenne temporelle de φ au cours de l'évolution en partant de la condition initiale x converge vers la moyenne spatiale de φ, à savoir l'intégrale de φ sur $[0, 1]$. On peut imaginer en quelque sorte que cela signifie que lors de son évolution temporelle en partant de x, le système va visiter toutes les régions de l'intervalle $[0, 1]$ de manière égale. Cela semble faux, et l'est, sans hypothèse supplémentaire sur la transformation T. On pourrait en effet imaginer qu'il existe une partie mesurable $E \subset [0, 1]$ qui soit invariante par T, à savoir $T^{-1}(E) = E$ et de mesure non triviale : $0 < |E| < 1$. Alors la mesure de Lebesgue restreinte à E est

invariante par T, au sens donné dans la définition ci-dessus, et le théorème de Birkhoff, s'il était valable, nous dirait que les moyennes de Birkhoff convergent presque toutes vers la moyenne de φ sur E, qui peut différer de la moyenne de φ sur $[0,1]$. Cela nous amène à la définition suivante :

Définition 11.11 — *On dit que la transformation préservant la mesure de Lebesgue* $\mathrm{T} : [0,1] \longrightarrow [0,1]$ *est* ergodique *si tout ensemble mesurable* E *tel que* $\mathrm{T}^{-1}(\mathrm{E}) = \mathrm{E}$ *satisfait à* $|\mathrm{E}| = 0$ *ou* $|\mathrm{E}| = 1$.

Dit autrement, la transformation est ergodique si l'on ne peut casser l'intervalle $[0,1]$ en bouts plus petits sur lesquels la mesure de Lebesgue est encore invariante. C'est une notion de minimalité ou d'irréductibilité comme on en trouve de nombreuses en mathématiques. Nous sommes en tout cas maintenant en mesure d'énoncer le théorème ergodique :

Théorème 11.31 — *Soit* $\mathrm{T} : [0,1] \longrightarrow [0,1]$ *une application mesurable, préservant la mesure de Lebesgue et ergodique ; soit encore* $\varphi \in \mathrm{L}^1([0,1])$ *une fonction Lebesgue intégrable sur l'intervalle. Alors pour presque tout* $x \in [0,1]$ *on a*

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(\mathrm{T}^k(x)) = \int_0^1 \varphi(x)\,\mathrm{d}x.$$

Il existe de nombreuses preuves de ce théorème. Toutes ou presque ont en commun de nécessiter la preuve d'une inégalité dite *maximale*. Ces inégalités se retrouvent dans de nombreuses branches de l'analyse et montrent en général une inégalité entre des moyennes locales d'une fonction ou de ses réarrangements et la moyenne globale de f. Un exemple frappant est donné par l'inégalité maximale de Hardy et Littlewood, utile pour démontrer le théorème des nombres premiers (celui qui dit que la fonction $\pi(x)$ définie sur $\mathbf{R}^+$ par $\pi(x) = \mathrm{Card}\{p \in \mathbf{N}, p \text{ premier et } p \leqslant x\}$ est équivalente, quand x tend vers l'infini, à $x/\operatorname{Log} x$). La démonstration de l'inégalité de Hardy et Littlewood fait l'objet d'un problème donné dans le chapitre consacré aux exercices.

Revenons à nos moutons et commençons par un petit lemme dont nous laissons la démonstration en exercice pour le lecteur :

Lemme 11.32 — *Avec les notations du théorème, démontrer que pour tout* $k \geqslant 0$ *on a*

$$\int_0^1 \varphi(\mathrm{T}^k(x))\,\mathrm{d}x = \int_0^1 \varphi(x)\,\mathrm{d}x.$$

Comme indication pour la preuve, nous conseillons au lecteur de penser aux fonctions en escalier et aux théorèmes de densité. Énonçons maintenant l'inégalité maximale :

Lemme 11.33 — *Avec les notations du théorème, notons*

$$\varphi^*(x) = \sup_{n \in \mathbf{N} - \{0\}} \sum_{k=0}^{n-1} \varphi(\mathrm{T}^k(x))$$

et $E^* = \{x \in [0,1], \varphi^*(x) \geqslant 0\}$*. Alors on a l'inégalité*

$$\int_{E^*} \varphi(x)\,dx \geqslant 0.$$

Avant de nous lancer tête baissée dans la démonstration, il peut être intéressant d'analyser l'énoncé. Si $x \in [0,1]$ est une condition initiale de notre dynamique, il va évoluer au cours du temps et visiter $T(x)$, $T^2(x)$, etc. On peut alors s'intéresser aux moyennes successives de la fonction φ le long de cette orbite entre le temps 0 et le temps n, pour n variant de 0 à l'infini. Si l'on se donne un temps n, on construit donc, comme nous l'avons vu ci-dessus, la moyenne de Birkhoff de φ le long de l'orbite de x entre le temps 0 et le temps n, que nous avons notée $S_n(\varphi)(x)$. La suite des moyennes $(S_n(\varphi)(x))$ est en général non constante et va présenter des oscillations qui, comme l'affirme le théorème ergodique, vont aller décroissant autour de la valeur limite donnée par la moyenne de φ sur $[0,1]$. Quitte à rajouter une constante à la fonction φ, on peut se limiter à essayer de comprendre ces oscillations autour de la valeur 0 et quitte même à changer φ en $-\varphi$ on peut se contenter de mesurer ce qui se passe pour les oscillations de cette moyenne *au-dessus* de la valeur 0. Ainsi la partie E^* définie dans le lemme n'est jamais que l'ensemble de x pour lesquels *une* moyenne de Birkhoff au moins prend une valeur positive. Dit autrement x est dans E^* si lorsque l'on enregistre les valeurs successives $S_0(\varphi)(x) = \varphi(x)$, $S_1(\varphi)(x) = (\varphi(x) + \varphi(T(x)))/2\ldots$ on observe à un moment donné une valeur positive. Bien évidemment E^* contient la partie de $[0,1]$ sur laquelle $\varphi(x) \geqslant 0$. Mais comme on fait des moyennes sur des temps arbitrairement longs il se peut que E^* contienne aussi beaucoup de points en lesquels φ est négative, valeurs négatives qui seront compensées dans le futur par l'apparition de suffisamment de valeurs positives pour que $S_n(\varphi)(x) \geqslant 0$. Le lemme maximal affirme cependant que φ est, en moyenne, positive sur E^*.

Démonstration — Remarquons tout d'abord que φ^* est bien une fonction mesurable et donc que E^* est une partie mesurable de $[0,1]$.

Afin de simplifier la démonstration et de nous concentrer sur les idées importantes, nous allons nous contenter de démontrer l'inégalité dans le cas où φ est bornée sur $[0,1]$ par une constante C. Pour $i \geqslant 1$ on pose

$$F_i = \left\{x,\ \sum_{k=0}^{i-1} \varphi(T^k(x)) \geqslant 0\right\}.$$

Si, pour $n \geqslant 1$ on note E_n l'union de tous les $F_i, i = 1,\ldots,n$, alors on a bien évidemment

$$E^* = \bigcup_{n\geqslant 1} E_n.$$

Comme φ est positive sur E^* et que l'union ci-dessus est une union croissante, en écrivant les intégrales sous la forme

$$\int_{E_n} \varphi(x)\,dx = \int_0^1 \varphi(x)\mathbf{1}_{E_n}(x)\,dx$$

le théorème de convergence monotone nous assure qu'il suffit de démontrer que l'on a, pour tout $n \geqslant 1$, l'inégalité

$$\int_{\mathrm{E}_n} \varphi(x)\,\mathrm{d}x \geqslant 0.$$

Fixons $n \geqslant 1$ et notons pour simplifier $\mathrm{E}_n = \mathrm{E}$. Soit $l > n$ un entier que nous ferons par la suite tendre vers l'infini. Réécrivons l'intégrale ci-dessus en utilisant le fait que T préserve la mesure :

$$\int_{\mathrm{E}} \varphi(x)\,\mathrm{d}x = \int_0^1 \varphi(x)\mathbf{1}_{\mathrm{E}}(x)\,\mathrm{d}x = \int_0^1 \frac{1}{l+1}\sum_{k=0}^{l} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x))\,\mathrm{d}x.$$

Nous allons démontrer que pour tout $l > n$ et tout $x \in [0,1]$ on a l'inégalité

$$(\dagger) \qquad \frac{1}{l+1}\sum_{k=0}^{l} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x)) \geqslant \frac{-n}{l+1}\mathrm{C},$$

ce qui, en faisant tendre l vers l'infini, démontrera le lemme pour φ bornée par C. Dans l'inégalité ci-dessus apparaît le terme $\varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x))$ qui vaut 0 si $\mathrm{T}^k(x) \notin \mathrm{E}$ et $\varphi(\mathrm{T}^k(x))$ si $x \in \mathrm{E}$. On doit donc contrôler les visites de l'orbite de x sous l'action de T dans E entre le temps 0 et le temps l. Or E est défini comme l'union des F_i pour i entre 1 et n. Ainsi si $x \in \mathrm{F}_i$ la première partie de la somme ci-dessus, jusqu'à l'indice $i-1$, est positive ou nulle et ne nous gêne nullement. Mais $i \leqslant n < l$: il va nous falloir passer à un contrôle des termes de l'inégalité pour les temps $k \geqslant i$ et nous allons pour cela utiliser une technique très courante en dynamique, consistant à découper l'intervalle de temps $[0,l]$ qui nous intéresse en petits intervalles de longueurs $\leqslant n$ sur lesquelles on contrôlera bien les termes de la somme ci-dessus.

Pour chaque x nous allons « rogner » la grosse somme ci-dessus en partant du temps 0 pour arriver jusqu'au temps l en commençant par arrêter la sommation suffisamment tôt, en un temps $\tau(x)$, pour que l'on ait

$$\sum_{k=0}^{\tau(x)-1} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x)) \geqslant 0.$$

Si $x \notin \mathrm{E}$ alors $x \notin \mathrm{F}_1$, et donc $\varphi(x) < 0$. On en déduit que pour tout $x \in [0,1]$ on a $\varphi(x)\mathbf{1}_{\mathrm{E}}(x) \geqslant \varphi(x)$. Si $x \notin \mathrm{E}$ on pose $\tau(x) = 1$ et l'on a bien dans ce cas

$$\sum_{k=0}^{\tau(x)-1} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x)) = \varphi(x)\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x)) = 0 \geqslant 0.$$

Si $x \in \mathrm{E}$ on définit $\tau(x)$ comme étant le plus grand entier $i \leqslant n$ tel que $x \in \mathrm{F}_i$. Par définition de F_i on a alors

$$\sum_{k=0}^{\tau(x)-1} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}_{\mathrm{N}}}(\mathrm{T}^k(x)) \geqslant 0.$$

Pour résumer, nous avons défini une fonction $\tau(x) \in \{1, \ldots, n\}$ telle que pour tout $x \in [0,1]$ on ait

$$(\star) \qquad \sum_{k=0}^{\tau(x)-1} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E_N}}(\mathrm{T}^k(x)) \geqslant 0.$$

Revenons à la somme de l'inégalité (†) et rognons-là par sa gauche :

$$\sum_{k=0}^{l} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x)) = \sum_{k=0}^{\tau(x)-1} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(x) + \sum_{\tau(x)}^{l} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(x).$$

Le premier terme du second membre est positif et ne nous gêne pas dans la démonstration de l'inégalité (†), nous pouvons donc nous concentrer sur le second et le rogner à nouveau :

$$\begin{aligned}\sum_{\tau(x)}^{l} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(x) = & \sum_{k=\tau(x)}^{\tau(x)+\tau(\mathrm{T}^{\tau(x)}(x))-1} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x)) \\ & + \sum_{k=\tau(x)+\tau(\mathrm{T}^{\tau(x)}(x))}^{l} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x)).\end{aligned}$$

Le premier terme du second membre est positif ou nul car on peut l'écrire

$$\sum_{k=0}^{\tau(\mathrm{T}^{\tau(x)}(x))-1} \varphi\big(\mathrm{T}^k(\mathrm{T}^{\tau(x)}(x))\big)\mathbf{1}_{\mathrm{E}}\big(\mathrm{T}^k(\mathrm{T}^{\tau(x)}(x))\big)$$

qui est positif d'après l'inégalité ($\star$) appliquée en $\mathrm{T}^{\tau(x)}(x)$.

Comme $\tau(x) \geqslant 1$ ce procédé rogne de manière effective la somme de (†) ; du fait que $\tau(x) \leqslant n$ il faut nous méfier lorsque l'on arrive au premier indice pour lequel $\tau(x)+\tau(\mathrm{T}^{\tau(x)}(x))+\cdots$ est $\geqslant l-n$. Mais les termes qui restent dans cette somme sont en nombre inférieur à n et tous minorés par $-\mathrm{C}$. Nous avons donc bien démontré que

$$\sum_{k=0}^{l} \varphi(\mathrm{T}^k(x))\mathbf{1}_{\mathrm{E}}(\mathrm{T}^k(x)) \geqslant -n\mathrm{C},$$

ce qui achève la démonstration de l'inégalité (†) et donc du lemme dans le cas où φ est bornée. Le cas général se démontre par des arguments standards que nous laissons le lecteur assidu retrouver : on coupe φ en posant $\varphi_n(x) = \varphi(x)$ si $|\varphi(x)| \leqslant n$ et $\varphi_n(x) = n$ sinon, puis on utilise un théorème de convergence. Il faut cependant faire attention au fait que les parties E^* dans ces constructions dépendent de la fonction choisie... Pour finir, remarquons que la preuve ne fait appel qu'à l'invariance de la mesure de Lebesgue, pas à l'ergodicité. □

Nous pouvons maintenant démontrer le théorème ergodique et, comme dans le cas de l'inégalité maximale, nous allons nous contenter de le démontrer pour une fonction φ bornée par une constante C. Le théorème est trivialement vrai pour les fonctions constantes. Quitte à additionner à φ une telle fonction, on peut supposer que

$$\int_0^1 \varphi(x)\,\mathrm{d}x = 0.$$

Soit $\varepsilon > 0$ et posons

$$\mathrm{U}_\varepsilon = \big\{x \in [0,1],\ \limsup_{n\to+\infty} \mathrm{S}_n(\varphi)(x) > \varepsilon\big\}$$

et

$$\mathrm{V}_\varepsilon = \big\{x \in [0,1],\ \limsup_{n\to+\infty} \mathrm{S}_n(\varphi)(x) < -\varepsilon\big\}.$$

Il nous suffit de démontrer que pour tout $\varepsilon > 0$ on a $|\mathrm{U}_\varepsilon| = |\mathrm{V}_\varepsilon| = 0$. En effet, si l'on pose

$$\mathrm{U} = \bigcup_{n\geqslant 1} \mathrm{U}_{1/n} \quad \text{et} \quad \mathrm{V} = \bigcup_{n\geqslant 1} \mathrm{V}_{1/n},$$

alors $\mathrm{U}\cup\mathrm{V}$ est de mesure nulle et pour tout $x \notin \mathrm{U}\cup\mathrm{V}$ la moyenne de Birkhoff $\mathrm{S}_n(\varphi)(x)$ converge vers 0, ce qui est le résultat recherché.

Raisonnons par l'absurde et supposons qu'il existe $\varepsilon > 0$ tel que $|\mathrm{U}_\varepsilon| > 0$ ou $|\mathrm{V}_\varepsilon| > 0$. Quitte à changer φ en $-\varphi$, supposons que l'on a $|\mathrm{U}_\varepsilon| > 0$. L'ensemble U_ε est invariant par T et comme T est ergodique on a nécessairement $|\mathrm{U}_\varepsilon| = 1$. Posons alors $\psi = \varphi - \varepsilon$ et

$$\mathrm{F}_i = \left\{x,\ \sum_{k=0}^{i-1} \psi(\mathrm{T}^k(x)) \geqslant 0\right\}.$$

Si $x \in \mathrm{U}$ alors il y a une infinité d'indices i pour lesquels $\mathrm{S}_i(\varphi)(x) > \varepsilon$ et donc

$$\mathrm{U} \subset \bigcup_{i\geqslant 1} \mathrm{F}_i.$$

Le second membre de cette inclusion n'est autre que l'ensemble E^* du lemme maximal appliqué à la fonction ψ. On a donc $|\mathrm{E}^*| = 1$ et ainsi

$$0 = \int_0^1 \varphi(x)\,\mathrm{d}x = \int_{\mathrm{E}^*} \varphi(x)\,\mathrm{d}x = \int_{\mathrm{E}^*} \psi(x)\,\mathrm{d}x + \varepsilon \geqslant \varepsilon,$$

ce qui est une contradiction. Nous avons donc démontré le théorème ergodique de Birkhoff.

Cela est bel et bon, mais s'il est facile de montrer que la transformation [10] : $[0,1] \longrightarrow [0,1]$ définie plus haut préserve la mesure de Lebesgue, il semble plus difficile de montrer qu'elle est ergodique, ce qui est indispensable si l'on veut lui appliquer le théorème de Birkhoff. Comment diable pouvons-nous nous en sortir ?

Il y a évidemment plusieurs manières de procéder et nous allons en présenter deux, de natures assez différentes. La première est une application immédiate de la définition de l'ergodicité. Soit E un ensemble mesurable et [10]-invariant. Notre but étant de démontrer que $|\mathrm{E}| = 0$ ou 1, on va supposer que $|\mathrm{E}| > 0$ et prouver qu'alors $|\mathrm{E}| = 1$. Fixons un $\varepsilon > 0$. Alors pour tout $\delta > 0$ assez petit il existe un intervalle $\mathrm{I} \subset [0, 1]$ tel que $|\mathrm{I}| = \delta$ et $|\mathrm{I} \cap \mathrm{E}| > (1 - \varepsilon)\delta$, sinon E serait de mesure nulle. Prenons maintenant n suffisamment grand et trouvons un intervalle I tel-que $|\mathrm{I}| = 10^{-n}$ et $|\mathrm{I} \cap \mathrm{E}| > (1 - \varepsilon)m^{-n}$. L'application [10] multiplie localement les longueurs d'un facteur 10. On a donc $[10]^n|\mathrm{I}| = 10^n.10^{-n} = 1$. Comme $\mathrm{T}^{-1}(\mathrm{E}) = \mathrm{E}$, on a aussi $\mathrm{T}(\mathrm{E}) = \mathrm{E}$; comme T préserve la mesure de Lebesgue on en déduit que

$$|\mathrm{E}| \geqslant |\mathrm{I} \cap \mathrm{E}| = |\mathrm{T}^n(\mathrm{I}) \cap \mathrm{E}| > 10^n(1 - \varepsilon)10^{-n} = (1 - \varepsilon).$$

Il s'ensuit alors que $|\mathrm{E}| > 1 - \varepsilon$, et ε ayant été pris arbitrairement petit, on a bien $|\mathrm{E}| = 1$.

Une deuxième approche consiste à remarquer que la transformation [10] n'est jamais que la multiplication par 10 modulo 1 sur $\mathbf{R}$ et qu'elle provient donc du morphisme du groupe unité $\mathbf{U}$ de $\mathbf{C}$ donné par $z \mapsto z^{10}$. Là où apparaît le groupe $\mathbf{U}$ apparaissent en même temps les séries de Fourier. Les séries de Fourier traitent de fonctions et non de parties mesurables, aussi aurons-nous besoin de la caractérisation suivante de l'ergodicité :

Proposition 11.34 — *La transformation préservant la mesure de Lebesgue* $\mathrm{T} : [0, 1] \longrightarrow [0, 1]$ *est ergodique si et seulement si toute fonction* φ *de classe* $\mathcal{C}^k$ *qui est* T*-invariante (c.-à-d.* $\varphi \circ \mathrm{T} = \varphi$*) est constante.*

Nous laissons la démonstration de cette proposition au lecteur et donnerons pour indication qu'un théorème de densité comme vu plus haut est un outil pratique pour passer d'un énoncé sur des parties mesurables à un énoncé sur des fonctions L^1. Revenons à l'ergodicité de [10] et considérons une fonction [10]-invariante φ de classe $\mathcal{C}^1$. On peut la prolonger en une fonction 1-périodique et la développer en série de Fourier :

$$\varphi(x) = \sum_{k=-\infty}^{+\infty} c_k \mathrm{e}^{2i\pi kx},$$

où les c_k sont les coefficients de Fourier de f. De l'identité $\varphi \circ [10] = \varphi$ on tire que pour tout $k \in \mathbf{Z}$ on a $c_k = c_{10k}$. Mais la fonction φ étant Lebesgue-intégrable sur $[0, 1]$, ses coefficients des Fourier tendent vers 0 quand $|k|$ tend vers l'infini, d'après le lemme de Lebesgue. Il s'ensuit que pour tout $k \neq 0$ on a $c_k = 0$ et donc que φ est constante.

On peut donner d'autres preuves de l'ergodicité, en remarquant par exemple que si $x \in [0, 1]$, alors ses 10^n préimages $[10]^{-n}(x) = \{y \in [0, 1],\ [10]^n y = x\}$ se distribuent de manière uniforme et harmonieuse dans l'intervalle $[0, 1]$. Ainsi si φ est [10]-invariante et *continue*, et si $x \in [0, 1]$, alors pour tout $n \geqslant 1$ et tout $y \in [10]^{-n}(x)$ on a $\varphi(y) = \varphi(x)$. Mais les 10^n points $y \in [10]^{-n}(x)$ se répartissent équitablement dans $[0, 1]$ et la continuité de φ entraîne alors que cette dernière doit être constante.

Maintenant que l'on sait que [10] est ergodique, à quoi cela peut-il nous servir ? La multiplication par 10 modulo 1 est intimement reliée à la décomposition d'un nombre réel en base 10. Pour tout $x \in [0,1]$ il existe une suite de chiffres $(a_i)_{i\geqslant 1}$ avec $a_i \in \{0,1,2,\ldots,9\}$ telle que

$$x = \sum_{i=1}^{+\infty} a_i 10^{-i}.$$

On écrit alors, comme en notation décimale classique : $x = 0, a_1 a_2 a_3 \ldots$. La suite (a_i) est unique sauf pour les x qui sont décimaux (c.-à-d. de la forme $k.10^{-n}$ pour certains $k, n \in \mathbf{N}$) pour lesquels on possède deux développements. Ainsi par exemple a-t-on

$$\frac{1}{10} = 0{,}1000\ldots = 0{,}0999\ldots$$

L'ensemble des décimaux est dénombrable et donc de mesure nulle. Nous les ignorerons dans la suite. Si l'on applique [10] au nombre $x = 0, a_1 \ldots$ on obtient le nombre $x' = 0, a_2 a_3 \ldots$ obtenu en décalant la suite des chiffres a_i vers la gauche et en effaçant le premier a_1. Ainsi itérer [10] revient à explorer les décimales successives du nombre x. Soit $c \in \{1,2,\ldots,9\}$ et $\varphi_c : [0,1] \longrightarrow [0,10]$ la fonction indicatrice de la partie $\{a_1 = c\}$. C'est une fonction L^1 et on peut lui appliquer le théorème ergodique de Birkhoff. Celui-ci nous dit alors que pour presque tout $x \in [0,1]$ la fréquence d'apparition du chiffre c dans le développement de x est égal, asymptotiquement, à $1/10$. Ainsi en prenant un nombre réel au hasard tombe-t-on sur un nombre dont le développement décimal contient autant de 0 que de 1 que de 9 ou que de 7. Ce raisonnement peut se généraliser au calcul de fréquence des occurrences des 10^n suites possibles de chiffres successifs et montre ainsi que presque tout nombre x est ce que l'on appelle un nombre normal en base 10. Il y a donc beaucoup de tels nombres, mais en construire un de manière explicite est une autre paire de manches, et nous vous invitons à vous les retrousser dans l'instant !

Il est cependant souvent plus intéressant d'étudier les nombres réels non *via* leur développement décimal mais au travers de leur développement en fraction continue. Et c'est alors la transformation de Gauß $\mathrm{G}(x) = \{x^{-1}\}$ qui intervient. Cette dernière ne préserve pas la mesure de Lebesgue mais on peut démontrer qu'il existe une fonction $\rho : [0,1] \longrightarrow \mathbf{R}$ qui est continue, partout strictement positive, d'intégrale sur $[0,1]$ égale à 1, et qui satisfait à

$$\forall \varphi \in \mathrm{L}^1([0,1]), \quad \int_0^1 \varphi \circ \mathrm{G}(x)\rho(x)\,\mathrm{d}x = \int_0^1 \varphi(x)\rho(x)\,\mathrm{d}x.$$

On dit que G préserve la mesure à densité $\rho(x)\,\mathrm{d}x$. Le théorème ergodique de Birkhoff s'étend sans peine à cette situation et même plus généralement à tout système préservant une *mesure de probabilité abstraite*. On entre alors dans le monde de la théorie abstraite de la mesure, des probabilités et des processus stochastiques, où le théorème ergodique prend le nom de loi des grands nombres.

Partie IV

Exercices, fascicule de résultats

This page intentionally left blank

Chapitre 12
Exercices

Dans cet avant-dernier chapitre, nous avons rassemblé les énoncés de quelques exercices et problèmes en rapport avec le cours. La plupart ne présentent aucune originalité. Les exercices sans mention particulière de difficulté sont considérés comme faisables par quiconque prétend avoir un tant soit peu assimilé le cours. Certains sont d'une difficulté supérieure et cela est signalé par une astérisque. Les exercices sont rassemblés par chapitre même si cette classification est parfois inadéquate.

12.1. Chapitre 1

Exercice 1

Soient $r \in \mathbf{Q}$ et $x \notin \mathbf{Q}$.

1. Démontrer que $r + x \notin \mathbf{Q}$, et que si de plus $r \neq 0$, alors $rx \notin \mathbf{Q}$.
2. Démontrer que pour tous nombres réels $a < b$, il existe des entiers p, q tels que $a < p/q < b$ (penser au fait que $\mathbf{R}$ est archimédien).
3. Démontrer que $\sqrt{2} \notin \mathbf{Q}$.
4. En déduire qu'entre deux nombres rationnels distincts il y a toujours un nombre irrationnel.

Exercice 2

Décrire un algorithme (utilisant des multiplications par 10, des divisions euclidiennes ou la fonction « partie entière ») permettant de trouver le développement décimal d'un nombre rationnel p/q. L'appliquer à $7/9, 29/64, 17/15$.

Exercice 3

Soient f, g, h des fonctions définies sur un intervalle I et à valeurs dans $\mathbf{R}$. On définit respectivement :

– le maximum de f et g, noté $f \vee g$ ou $\max\{f, g\}$, par

$$\max\{f,g\}(x) = \max\big\{f(x), g(x)\big\}\;;$$

– le minimum de f et g, noté $f \wedge g$ ou $\min\{f, g\}$, par

$$\min\{f, g\}(x) = \min\big\{f(x), g(x)\big\}\,;$$

– les parties positives et négatives de f, notées f^+ et f^-, par

$$f^+ = f \vee 0,\ f^- = (-f) \vee 0\,;$$

– la médiane de f, g et h, notée $\text{méd}\{f, g, h\}$, par le fait que $\text{méd}\{f, g, h\}(x)$ est celui des trois nombres $f(x), g(x), h(x)$ situé entre les deux autres.

1. Sur $\mathrm{I} = [0, 2\pi]$, on pose $f(x) = \sin x$, $g(x) = \cos x$, et $h(x) = (2\pi)^{-1}x$. Tracer les graphes de f, g, h ainsi que de toutes les fonctions définies à partir de f, g, h ci-dessus.
2. Montrer les égalités suivantes :
   (a) $f \vee g = \frac{1}{2}(f + g + |f - g|)$ et $f \wedge g = \frac{1}{2}(f + g - |f - g|)\,;$
   (b) $f^+ = \frac{1}{2}(f + |f|)$ et $f^- = \frac{1}{2}(|f| - f)\,;$
   (c) $f = f^+ - f^-$ et $|f| = f^+ + f^-\,;$
   (d) $\text{méd}\{f, g, h\} = \max\{\min\{f, g\}, \min\{g, h\}, \min\{h, f\}\}\,;$
   (e) une égalité similaire à la précédente mais commençant par un min ;
   (f) $\min\{a, b, c\} = \min\{\min\{a, b\}, c\}$.

Exercice 4

Soient f et g deux fonctions numériques définies sur $\mathbf{R}^+$.

1. Écrire et expliquer la définition de $\lim_{x\to+\infty} f(x) = l$, $l \in \mathbf{R}$.
2. On suppose ici que
$$g > 0 \quad \text{et} \quad \lim_{x\to+\infty} \frac{f(x)}{g(x)} = l \neq 0.$$
   (a) Montrer que $\lim_{x\to+\infty} f(x) = 0$ si, et seulement si, $\lim_{x\to+\infty} g(x) = 0$.
   (b) Si $l > 0$, montrer que $\lim_{x\to+\infty} f(x) = +\infty$ si, et seulement si $\lim_{x\to+\infty} g(x) = +\infty$.
3. Montrer que si f est périodique et non constante, alors elle n'admet pas de limite en $+\infty$.
4. On suppose maintenant que f est positive et croissante, et que $\lim_{x\to+\infty}(f(x+1) - f(x)) = 0$. Montrer que
$$\lim_{x\to+\infty} \frac{f(x)}{x} = 0.$$

Exercice 5

Soient K une partie fermée et bornée de $\mathbf{R}$ et $[a, b]$ un segment contenant K. En utilisant le théorème des segments emboîtés et un découpage successif du segment $[a, b]$ en deux parties égales, démontrer par l'absurde le théorème de Heine-Borel : de $\mathcal{U}$ on peut extraire un sous-recouvrement fini.

Exercice 6 (*)

Soit f une fonction dérivable sur l'intervalle $[a, b]$, admettant en a et b respectivement des dérivées à droite et à gauche.

1. Rappeler l'énoncé de la formule des accroissements finis. On notera c l'un des points de $]a, b[$ donnés par celle-ci.
2. Soit f la fonction définie sur $\mathbf{R}$ par
$$f(x) = \alpha + \beta x + \eta e^{\zeta x},$$
où α, β sont quelconques et η, ζ non nuls.
   (a) Déterminer, en fonction de X, le point c donné par la formule des accroissements finis appliquée à f sur l'intervalle $[0, \mathrm{X}]$.
   (b) En déduire que la fonction
$$h(x) = \frac{1}{\zeta x} \operatorname{Log} \frac{e^{\zeta x} - 1}{\zeta x}$$
   est bornée sur $\mathbf{R}$.
3. On suppose ici que f est définie et dérivable sur $[a, +\infty[$ et que $\lim_{x\to+\infty} f(x) = f(a)$. Montrer qu'il existe $c \in]a, +\infty[$ tel que $f'(c) = 0$.
4. En appliquant le théorème des accroissements finis à la fonction logarithme sur $[n, n+1]$, montrer que
$$\mathrm{S}_n = \sum_{k=1}^{n} \frac{1}{k}$$
tend vers l'infini quand n tend vers l'infini.

Exercice 7

Soit $0 < \alpha \leqslant 1$. Montrer que pour tout $t \geqslant 0$ on a
$$t^\alpha \leqslant \alpha t + (1 - \alpha),$$
et que l'égalité n'a lieu que pour $t = 1$.

Exercice 8

Soient $p \in]1, +\infty[$ et q tel que $p^{-1} + q^{-1} = 1$. Montrer que pour tous $\mathrm{A}, \mathrm{B} \in \mathbf{R}^+$, on a
$$\mathrm{AB} \leqslant \frac{1}{p}\mathrm{A}^p + \frac{1}{q}\mathrm{B}^q.$$

Exercice 9

Soit (U_n) une suite de nombres (réels ou complexes, voire même d'éléments d'un anneau quelconque). La suite (u_n) définie par $u_n = \mathrm{U}_{n+1} - \mathrm{U}_n$ s'appelle la dérivée discrète de la suite (U_n), et (U_n) est une *primitive discrète* de la suite (u_n).

1. Montrer que toute suite (u_n) admet une primitive discrète, et que celle-ci est unique à l'addition d'une constante près.
2. Soit (U_n) une primitive discrète de la suite (u_n). Déterminer la somme suivante :

$$\sum_{k=0}^{n} u_n.$$

La formule fondamentale du calcul différentiel et intégral peut donc être vue comme une version continue de la formule des sommes télescopiques.

Exercice 10 (*)

Soit $\mathrm{I} = [a, b]$ un intervalle compact. On dit que $f : [a, b] \longrightarrow \mathbf{R}$ est réglée si elle est limite uniforme d'une suite de fonctions en escalier.

1. Montrer que f est réglée si, et seulement si, pour tout $\varepsilon > 0$ il existe une fonction en escalier ψ_ε telle que

$$|f(x) - \psi_\varepsilon(x)| \leqslant \varepsilon, \ \ \forall x \in [a, b].$$

2. Montrer qu'une fonction continue est réglée.
3. Soit $f : [0, 1] \longrightarrow \mathbf{R}$ la fonction définie par

$$f(0) = 0, \quad f(x) = 1/\sqrt{x} \ \text{ si } \ x \neq 0.$$

Montrer que f n'est pas réglée.
4. Montrer que f est réglée si, et seulement si, elle admet en tout point de $[a, b]$ des limites à droite et à gauche.
5. Montrer qu'une fonction monotone est réglée.
6. Montrer qu'une fonction réglée est bornée.
7. Montrer que l'ensemble des points de discontinuité d'une fonction réglée est dénombrable.

Exercice 11 (*)

Soit f une fonction monotone sur un intervalle compact $[a, b]$ de $\mathbf{R}$.

1. Montrer que f admet une limite à droite et une limite à gauche en tout point $x \in [a, b]$. On notera ces limites respectivement $f(x^+)$ et $f(x^-)$.
2. Soit $a < x_1 < x_2 < \ldots < x_n < b$ une suite de points de $]a, b[$. Montrer que l'on a

$$\sum_{k=1}^{n} |f(x_k^+) - f(x_k^-)| \leqslant |f(b) - f(a)|.$$

3. Montrer que pour tout entier $n > 0$, l'ensemble des éléments $x \in \,]a, b[$ tels que

$$|f(x^+) - f(x^-)| > \frac{1}{n}$$

est fini.
4. En déduire que l'ensemble des points de discontinuité de f est dénombrable.

Problème : construction *ab initio* des fonctions exponentielles

Le but de ce devoir est de donner une construction directe des fonctions exponentielles. Si $a > 0$, on peut bien évidemment définir, pour tout entier $n \in \mathbf{N}$, le nombre a^n, et l'on a clairement $a^{n+m} = a^n a^m$ de sorte qu'en passant de n à a^n, on transforme les multiplications en additions. Nous allons voir que l'on peut, pour tout nombre réel x, définir a^x, donnant ainsi naissance à une fonction définie sur tout $\mathbf{R}$ par $x \mapsto a^x$ et qui s'appelle fonction exponentielle (de base a). Nous verrons plus tard (lorsque nous disposerons de la notion de dérivée) qu'il existe, parmi les fonctions exponentielles, une fonction plus jolie que les autres (une exponentielle canonique en quelque sorte) : « la » fonction exponentielle $x \mapsto \mathrm{e}^x$.

L'idée de la construction est simple et très commune : on définit l'exponentielle sur les nombres entiers, puis les nombres rationnels et l'on étend enfin le tout aux nombres réels en utilisant la densité de l'ensemble des nombres rationnels (soit donc l'axiome d'Archimède) et la complétude de $\mathbf{R}$ (à savoir la propriété de la borne supérieure).

1. Exponentielle des entiers

On se donne ici deux nombres réels a et b tous deux *strictement positifs.* On pose $a^0 = 1$ et, pour $n \geqslant 1$, on définit a^n par la relation de récurrence $a^n = a.a^{n-1}$, de sorte que $a^2 = a.a$, $a^3 = a.a.a$, etc. Cette partie est très simple et vous est, pour la plupart, connue. Nous vous demandons cependant de la rédiger avec soin.

1. Démontrer que pour tous $n, m \in \mathbf{N}$ on a
   - $a^n a^m = a^{n+m}$;
   - $(ab)^n = a^n b^n$;
   - $(a^n)^m = a^{nm}$;
   - $(a/b)^n = (a^n)/(b^n)$;
   - pour tout $n \geqslant m$, $(a^n)/(a^m) = a^{n-m}$.
   - si $a > 1$, alors $n < m$ si et seulement si $a^n < a^m$;
   - $a^n > 0$;
   - $a < b$ si et seulement si, pour un certain $n > 0$ on a $a^n < b^n$.

   Il pourra être parfois utile de recourir à la récurrence.
2. Pour $n \in \mathbf{N}$, on pose $a^{-n} = 1/a^n$, ce qui permet d'étendre la définition de a^n à tous les entiers relatifs. Montrer que les propriétés démontrées dans la question précédente s'étendent, sans restriction, aux entiers relatifs.

2. Exponentielle des inverses d'entiers

Nous aimerions définir $a^{1/n}$ pour $n \geqslant 1$. Si une telle écriture doit avoir un sens, alors l'on devrait avoir $(a^{1/n})^n = a^{n/n} = a$ et donc $a^{1/n}$ doit être une *racine* n*-ème* de a. Ne disposant pas d'outils avancés d'analyse comme le théorème des valeurs intermédiaires, nous allons montrer qu'une telle racine existe en utilisant la propriété de la borne supérieure. Dans toute cette partie, a est un nombre réel strictement positif.

1. Rappeler la formule du binôme.
2. Montrer qu'il existe une constante $\mathrm{K} > 0$ telle que, pour tout $0 \leqslant \varepsilon < 1$, on a
$$(1+\varepsilon)^n \leqslant 1 + \mathrm{K}\varepsilon.$$
3. Soit E une partie majorée de $\mathbf{R}$. Montrer que les deux propriétés suivantes sont équivalentes :
   - $\mathrm{S} = \sup \mathrm{E}$;
   - S est un majorant de E et pour tout $\varepsilon > 0$, il existe $z \in \mathrm{E}$ tel que $\mathrm{S} - \varepsilon < z \leqslant \mathrm{S}$.
4. On définit la partie suivante de $\mathbf{R}$:
$$\mathrm{E} = \big\{x \in \mathbf{R}, 0 \leqslant x^n < a\big\}.$$
Montrer que E est majorée (on distinguera les cas $a > 1$ et $a \leqslant 1$). Nous noterons S la borne supérieure de E.
5. Soit z tel que $z < a$. Montrer qu'il existe $\varepsilon > 0$ tel que $z(1 + \mathrm{K}\varepsilon) < a$.
6. Montrer par l'absurde que $\mathrm{S}^n \geqslant a$.
7. Montrer par l'absurde que $\mathrm{S}^n \leqslant a$. En conclure que $\mathrm{S}^n = a$.
8. Déduire de ce qui précède (et des propriétés de l'exponentielle pour les entiers démontrées dans la partie précédente) qu'il existe un unique nombre réel $z > 0$ tel que $z^n = a$ (et que ce nombre n'est autre que le nombre S défini ci-dessus).

Si maintenant $n > 1$ est un entier, nous noterons $a^{1/n}$ l'unique solution positive de l'équation $z^n = a$.

3. Exponentielle des rationnels

Nous voudrions dorénavant définir a^r lorsque r est un nombre rationnel. Un tel nombre peut s'écrire sous la forme $r = m/n$ avec m et n entiers et, de plus, $n > 0$. Le premier problème auquel nous sommes confrontés est qu'une telle écriture n'est pas unique : la définition que nous allons adopter devra assurer que $a^{2/3} = a^{4/6}$. Dans toute cette partie, a et b seront deux nombres réels strictement positifs.

1. Montrer que s'il existe un entier $p \geqslant 1$ tel que $a^p = b^p$, alors $a = b$.
2. Soient m, n, m', n' quatre entiers tels que n et n' soient strictement positifs et tels que $m/n = m'/n'$. Montrer que $(a^{1/m})^m = (a^{1/n'})^{m'}$.
3. En déduire que l'on peut bien définir a^r pour tout nombre rationnel $r \in \mathbf{Q}$.
4. Montrer maintenant que pour tous nombres réels $a, b > 0$ et tous nombres rationnels r, s, les propriétés suivantes sont vraies :
   - $a^r a^s = a^{r+s}$;
   - $(ab)^r = a^r b^r$;
   - $(a^r)^s = a^{rs}$;
   - $(a/b)^r = a^r/b^r$;
   - $a^r/a^s = a^{r-s}$;
   - si $a > 1$, alors $r < s$ si et seulement si $a^r < a^s$;

– $a^r > 0$;
– si $r > 0$, alors $a < b$ si et seulement si $a^r < b^r$.

4. Exponentielle des réels

Nous allons enfin pouvoir définir a^x pour tout nombre réel x. Cela ne se fera cependant pas sans mal. Les arguments vont s'enchaîner et il faudra savoir utiliser les résultats démontrés avec discernement. *Nous supposerons dans cette partie que* $a > 1$. Cette restriction peut être aisément retirée par la suite, nous en reparlerons.

1. Nous allons dans cette question montrer que la suite de terme général $u_n = a^{1/n}$ converge vers 1 quand n tend vers l'infini.
   (a) Démontrer que pour tous $x, y > 0$ on a
   $$\sqrt{x} - \sqrt{y} = \frac{x-y}{\sqrt{x}+\sqrt{y}}.$$
   (b) Déduire de ce qui précède, en utilisant la définition avec les *epsilon* et les *delta* que si (v_n) est une suite de nombres réels strictement positifs qui converge vers une limite $l > 0$, alors la suite $(\sqrt{(v_n)})$ converge vers $\sqrt{l}$.
   (c) Montrer que la suite (u_n) admet une limite quand n tend vers l'infini. Nous noterons l cette limite.
   (d) Montrer que l satisfait à $\sqrt{l} = l$. En déduire que $l = 1$.
2. Soit $x \in \mathbf{R}$. On pose
   $$\mathrm{E}_x = \{a^r,\ r \in \mathbf{Q} \ \text{ et } \ r < x\}.$$
   Montrer que E_x est majorée et non vide. En déduire qu'elle admet une borne supérieure, que nous noterons temporairement $\mathrm{S}(x)$. Nous allons maintenant démontrer que si $x \in \mathbf{Q}$ alors $\mathrm{S}(x) = a^x$.
   (a) Montrer que si $r \in \mathbf{Q}$ alors a^r est un majorant de E_r.
   (b) Montrer que pour tout $\varepsilon > 0$ il existe un entier N tel que, pour tout nombre rationnel $0 \leqslant s < 1/\mathrm{N}$ on a $|a^s - 1| < \varepsilon/a^r$. En déduire que $\mathrm{S}(r) = a^r$. **Nous noterons désormais $\mathrm{S}(x)$, pour $x \in \mathbf{R}$, sous la forme a^x.**
3. Nous allons démontrer que l'on a $a^x a^y = a^{x+y}$.
   (a) Soient A, B deux parties non vides de $\mathbf{R}$. On note AB la partie de $\mathbf{R}$ constituée des produits d'un élément de A et d'un élément de B, autrement-dit
   $$\mathrm{AB} = \{ab,\ a \in \mathrm{A},\, b \in \mathrm{B}\}.$$
   Déterminer AB pour $\mathrm{A} = [0,1]$ et $\mathrm{B} = [2,7]$.
   (b) Soient a, b deux nombres réels positifs. Montrer que pour tout $\varepsilon > 0$ il existe η satisfaisant aux conditions suivantes :
   $$0 < \eta < \frac{a+b}{2},\ \eta(a+b) - \eta^2 < \varepsilon \text{ et } ab - \varepsilon < (a-\eta)(b-\eta) < ab.$$
   (c) Montrer que si A, B sont deux parties non vides de $\mathbf{R}_+$, alors
   $$\sup(\mathrm{AB}) = \sup(\mathrm{A})\sup(\mathrm{B}).$$

(d) Soient x, y deux réels et X et Y les deux parties de $\mathbf{R}$ définies par

$$\mathrm{X} = \{r + s,\ r, s \in \mathbf{Q} \text{ et } r < x, s < y\}, \quad \mathrm{Y} = \{t \in \mathbf{Q},\ t < x + y\}.$$

Montrer que X = Y.

(e) Déduire de tout ce qui précède que pour tous réels x, y on a $a^{x+y} = a^x a^y$.

4. Nous allons maintenant démontrer que si $a, b > 1$ et $x \in \mathbf{R}$ alors $(ab)^x = a^x b^x$.

(a) Démontrer que

$$\sup\{a^r b^r, r \in \mathbf{Q}, r < x\} = \sup\{a^r b^s,\ r, s \in \mathbf{Q},\ r < x,\ s < x\}.$$

(b) En déduire que $(ab)^x = a^x b^x$.

5. Nous voulons maintenant montrer que si $a > 1$ et x, y sont strictement positifs, alors $(a^x)^y = a^{xy}$.

(a) Montrer que pour tout rationnel r et tout réel x tels que $0 < r < x$ on a $1 < a^r < a^x$.

(b) Soient A et B les deux parties définies par

$$\mathrm{A} = \{rs,\ 0 < r < x,\ 0 < s < y \text{ et } r, s \in \mathbf{Q}\} \text{ et } \mathrm{B} = \{t \in \mathbf{Q},\ 0 < t < xy\}.$$

Montrer que A = B.

(c) Montrer que

$$\sup\{(a^x)^s, 0 < s < y, s \in \mathbf{Q}\} = \sup\{(a^r)^s, r, s \in \mathbf{Q} \text{ et } 0 < r < x, 0 < s < y\}.$$

(d) Déduire de ce qui précède que $(a^x)^y = a^{xy}$.

6. Montrer que si (r_n) est une suite croissante de nombres rationnels tels que $r_n < x$ pour tout n et telle que $\lim_{n\to+\infty} r_n = x$, alors $\lim_{n\to+\infty} a^{r_n} = a^x$. On pourra raisonner par contradiction.

7. Montrer que si $a > b > 1$ et $x \in \mathbf{R}$ alors $(a/b)^x = a^x/b^x$.

8. Montrer que si $a > 1$ et $x > y$ alors $a^x/a^y = a^{x-y}$.

9. Montrer que si $a > 1$, alors $a^x < a^y$ si et seulement si $x < y$.

10. Montrer que $a^{-x} = 1/a^x$.

11. Soient $a, b > 1$ et $x > 0$. Montrer que $a < b$ si et seulement si $a^x < b^x$.

12. Montrer que si $a > 1$ alors $x \mapsto a^x$ est bijective. Sa bijection réciproque est notée $x \mapsto \operatorname{Log}_a x$ et s'appelle le *logarithme en base a*. Énoncer et démontrer les propriétés élémentaires et importantes du logarithme.

13. Montrer que l'on peut définir a^x pour $0 < a \leqslant 1$. Énoncer les propriétés des exponentielles en base < 1. Ces dernières sont peu utilisées.

5. Propriétés analytiques des exponentielles. Exponentielle naturelle

Nous allons ici démontrer que les fonctions exponentielles sont gentilles et que parmi elles il y en a une plus jolie que les autres. Pour simplifier les choses, nous fixerons ici a et supposerons qu'il est strictement plus grand que 1.

1. Soit $f : \mathbf{R} \longrightarrow \mathbf{R}$ une fonction strictement croissante et $x \in \mathbf{R}$. On suppose que pour toute suite strictement croissante de nombres rationnels (r_n) qui converge vers x on a $\lim_{n\to+\infty} f(r_n) = f(x)$ et que pour toute suite strictement décroissante de nombres rationnels (s_n) qui converge vers x on a $\lim_{n\to+\infty} f(s_n) = f(x)$. Montrer que f est continue en x.
2. Démontrer que la fonction exponentielle est continue sur $\mathbf{R}$.
3. Soit f une fonction continue sur $\mathbf{R}$ telle que, pour tous $x, y \in \mathbf{R}$ on ait $f(x+y) = f(x)f(y)$. On suppose que $f(1) = a > 0$.
   (a) Déterminer les valeurs de $f(n)$ pour $n \in \mathbf{Z}$. En déduire les valeurs de $f(r)$ pour $r = m/n$ rationnel.
   (b) En utilisant la densité des nombres rationnels dans $\mathbf{R}$, montrer que $f(x) = a^x$.
4. Nous allons ici montrer qu'il existe une exponentielle plus jolie que les autres, celle de base e.
   (a) On considère le problème suivant : trouver une fonction f définie sur $\mathbf{R}$ qui soit dérivable et satisfasse à $f' = f$ et $f(0) = 1$. Un tel problème s'appelle un *problème de Cauchy* consistant à trouver une solution de *l'équation différentielle* $y' = y$ avec *condition initiale* (ou *au bord*) $y(0) = 1$. La *méthode d'Euler* permet de construire des *solutions approchées* à de tels problèmes. Fixons $x > 0$ et n un entier positif très grand. Construisons sur $[0, x]$ la fonction affine par morceau f_n définie de la manière suivante. Sur l'intervalle $[0, x/n]$, la fonction f_n est affine et déterminée par $f_n(0) = 1$ et $f_n'(t) = f_n'(0) = 1$. De cette manière $f_n(x/n) = 1 + x/n$. Sur $[x/n, 2x/n]$, la fonction f_n est affine et déterminée par $f_n(x/n) = 1 + x/n$ et $f_n'(t) = f_n(x/n) = 1 + x/n$, de sorte que $f_n(2x/n) = (1 + x/n)^2$. Continuer ainsi la construction de la fonction f_n sur les intervalles $[2x/n, 3x/n], \ldots, [(n-1)x/n, x]$. Montrer que $f_n(x) = (1 + x/n)^n$. Nous allons dans la suite démontrer que la suite $(f_n(x))$ converge, pour tout x, vers un nombre que nous noterons $f(x)$ et que la fonction $x \mapsto f(x)$ est solution du problème de Cauchy décrit plus haut.
   (b) Soit $t > -1$ un nombre réel. Démontrer que pour tout entier positif n on a $(1+t)^n \geqslant 1 + nt$.
   (c) Soit $t < 1$ un nombre réel. Démontrer que pour tout entier positif n on a $(1-t)^n \geqslant 1 - nt$.
   (d) Montrer que pour $n \geqslant |x|$ la suite $f_n(x) = (1 + x/n)^n$ est monotone croissante. On pourra commencer par démontrer l'identité suivante :
   $$1 + \frac{x}{n+1} = \left(1 + \frac{x}{n}\right)\left(1 - \frac{x}{n(n+1)(1 + x/n)}\right).$$
   (e) Montrer de même que pour $n \geqslant |x|$ la suite $g_n(x) = (1 - x/n)^{-n}$ est décroissante.
   (f) Montrer que les suites $(f_n(x))$ et $(g_n(x))$ sont adjacentes ; on pourra pour cela s'intéresser au rapport $f_n(x)/g_n(x)$. *On note* $\exp(x)$ *la limite commune aux suites* $(f_n(x))$ *et* $(g_n(x))$.
   (g) Calculer la dérivée et la dérivée seconde de f_n. Quelle est la limite de $(f_n'(x))$ quand n tend vers l'infini ?
   (h) Montrer que, pour $n \geqslant |x| + 1$, la fonction f_n' est croissante sur $[x, x+1]$.

(i) Soit f une fonction dérivable sur un intervalle $[a,b]$ et l'on suppose qu'il existe deux constantes m, M telles que, pour tout $x \in [a,b]$ on ait $m \leqslant f'(x) \leqslant \mathrm{M}$. Montrer que pour tout $x \in [a,b]$ on a

$$m(x-a) \leqslant f(x) - f(a) \leqslant \mathrm{M}(x-a).$$

On pourra avec profit étudier une fonction comme $x \mapsto m(x-a)-(f(x)-f(a))$.

(j) Déduire de ce qui précède que

$$(1+h)\exp(x) \leqslant \exp(x+h) \leqslant \frac{\exp(x)}{1-h}.$$

(k) Montrer finalement que $x \mapsto \exp(x)$ est dérivable sur $\mathbf{R}$ et qu'elle est solution du problème de Cauchy $y' = y, y(0) = 1$.

(l) En utilisant la fonction $h(x) = \exp(x)\exp(-x)$, montrer que exp est *l'unique solution* du problème de Cauchy ci-dessus. En déduire en particulier que $\exp(-x) = 1/\exp(x)$.

(m) En fixant y et en étudiant la fonction $x \mapsto \exp(x+y) - \exp(x)\exp(y)$, montrer que la fonction exp satisfait à *l'équation fonctionnelle*

$$\exp(x+y) = \exp(x)\exp(y).$$

(n) Conclure que si l'on pose $\mathrm{e} = \exp(1)$, on a $\exp(x) = \mathrm{e}^x$ pour tout $x \in \mathbf{R}$.

Exercice 12

En utilisant la définition 6 du livre V des éléments d'Euclide, montrer que si a/b et a'/b' ne sont pas en même raison, il existe une proportion *rationnelle* qui peut s'insérer entre ces deux proportions.

12.2. Chapitres 2, 3 et 4

Exercice 1

Soit δ la jauge définie sur $[0,1]$ par

$$\delta(0) = \frac{1}{2},\ \delta(t) = \frac{1}{2}t \text{ si } t \neq 0.$$

Déterminer si les subdivisions suivantes sont ou non δ-fines :

$$\mathcal{P}_1 = \{([0,1/4],0),([1/4,1/2],1/2),([1/2,1],3/4)\},$$
$$\mathcal{P}_2 = \{([0,1/4],0),([1/4,1/2],1/2),([1/2,1],3/5)\},$$
$$\mathcal{P}_3 = \{([0,1/4],0),([1/4,1/2],1/2),([1/2,1],1/4)\}.$$

Calculer les sommes de Riemann $\mathrm{S}(f,\mathcal{P})$ pour les trois subdivisions ci-dessus et $f(x) = x^2$.

Exercice 2

Reprendre l'exercice précédent avec les jauges définies par

$$\delta(0) = \frac{1}{2}, \quad \delta(t) = \frac{3}{4}t \text{ si } t \neq 0$$

et

$$\delta(0) = \frac{1}{2}, \quad \delta(t) = \frac{9}{10}t \text{ si } t \neq 0$$

Exercice 3

Soit η la jauge définie sur $[0,1]$ par

$$\eta(0) = \eta(1) = 1/4, \quad \eta(t) = \min\left\{|t|, |1-t|\right\} \text{ si } t \in]0,1[.$$

1. Montrer que toute subdivision η-fine admet 0 et 1 comme points de marquage.
2. Construire une subdivision η-fine.

Exercice 4

Soit f la fonction nulle partout sur $]0,1]$ et égale à 1 en 0. Montrer que f est intégrable, et donner la valeur de son intégrale.

Exercice 5

Soit k la fonction définie sur $[0,1]$ par

$$k(x) = n \text{ si } x = 1/n, \; n \in \mathbf{N}^*, \quad k(x) = 0 \text{ sinon.}$$

Montrer que k est intégrable d'intégrale nulle.

Exercice 6

Montrer par récurrence sur sa taille qu'une combinaison linéaire de fonctions intégrables est intégrable.

Exercice 7

Montrer, en construisant une suite de fonctions en escalier qui la minorent, que la fonction

$$f(0) = 0, f(x) = 1/x, \; x \neq 0$$

n'est pas intégrable sur $[0,1]$.

Exercice 8

Soit $\mathrm{I} = [a,b]$ un intervalle compact et f une fonction continue sur I. Montrer que f est intégrable au sens de Riemann, au sens où l'on peut se restreindre aux jauges constantes. En déduire que si f est de signe constant et d'intégrale nulle, elle est nulle.

Exercice 9

Montrer que si f est continue sur $[a, b]$, alors il existe $c \in [a, b]$ tel que

$$\int_a^b f(x)\,\mathrm{d}x = (b-a)f(c).$$

Exercice 10

Soit f une fonction continue sur $\mathbf{R}$. On définit g sur $\mathbf{R}^*$ par

$$g(x) = \frac{1}{x}\int_0^x f(t)\,\mathrm{d}t.$$

1. Montrer que g se prolonge par continuité en 0.
2. Montrer que si f est périodique, alors $g(x)$ admet une limite quand x tend vers $+\infty$.

Exercice 11 (*)

Soit $(\mathrm{I}_k)_{k\in\mathbf{N}}$ une famille dénombrable d'intervalles ouverts et

$$\mathrm{E} = \bigcup_{k\in\mathbf{N}} \mathrm{I}_k$$

son union. Montrer qu'il existe $(\mathrm{J}_k)_k$ une famille dénombrable d'intervalles ouverts *deux à deux disjoints* telle que E soit égal à l'union des J_k. Montrer que si de plus on suppose que la somme des longueurs des I_k est majorée par un nombre M alors la somme des longueurs des J_k satisfait la même propriété.

Exercice 12

Montrer qu'une fonction monotone et bornée sur un intervalle compact $[a, b]$ est intégrable sur $[a, b]$.

Exercice 13

On rappelle qu'une fonction $f : [a, b] \longrightarrow \mathbf{R}$ est réglée si elle est limite uniforme sur $[a, b]$ d'une suite de fonctions en escaliers.

1. Montrer qu'une fonction réglée sur $[a, b]$ y est intégrable.
2. Soient f une fonction intégrable et minorée sur $[a, b]$ et g une fonction réglée sur $[a, b]$. Montrer que fg est intégrable sur $[a, b]$.
3. En déduire que si f est intégrable et minorée sur $[a, b]$ et que si g est continue (resp. monotone) sur $[a, b]$ alors fg est intégrable sur $[a, b]$.
4. Soient f une fonction positive et intégrable sur $[a, b]$ et g une fonction continue sur $[a, b]$. Montrer qu'il existe $c \in [a, b]$ tel que

$$\int_a^b f(x)g(x)\,\mathrm{d}x = g(c)\int_a^b f(x)\,\mathrm{d}x.$$

Ce résultat porte le nom de *formule de la moyenne.*

Exercice 14

Donner une valeur approchée de l'intégrale

$$\int_0^1 \frac{\mathrm{d}x}{1+x}$$

en utilisant la méthode de Simpson avec $n = 10$. Estimer l'erreur commise en utilisant le résultat général sur l'erreur dans la méthode de Simpson puis en calculant exactement l'intégrale.

12.3. Chapitres 5 et 6

Exercice 1

Calculer les primitives suivantes :

$$\int \mathrm{e}^x \cos x \,\mathrm{d}x, \quad \int \frac{\operatorname{Log} x}{x^n}\,\mathrm{d}x,\ n \in \mathbf{N}, \quad \int x \arctan x \,\mathrm{d}x, \quad \int (x^2+x+1)\mathrm{e}^x \,\mathrm{d}x.$$

Exercice 2

À l'aide de changements de variable, calculer les primitives suivantes :

$$\int \frac{1}{\sqrt{2+x}+\sqrt[3]{2+x}}\,\mathrm{d}x \quad \left(t = \sqrt[6]{2+x}\right),$$

$$\int \frac{1}{((x-1)^2-4)^2}\,\mathrm{d}x \quad \left(\operatorname{th} u = \frac{x-1}{2}\right),$$

$$\int (\arcsin x)^2 \,\mathrm{d}x, \quad \int x^2\sqrt{1+x^3}\,\mathrm{d}x.$$

Exercice 3

Calculer l'intégrale

$$\mathrm{I} = \int_0^{\operatorname{Log} 2} \sqrt{\mathrm{e}^x - 1}\,\mathrm{d}x.$$

Exercice 4

Nous vous proposons un exercice quelque peu pénible : en utilisant les formules donnant les racines des polynômes de degrés 1, 2, 3 et 4, écrire un programme informatique donnant la primitive d'une fraction rationnelle P/Q telle que $\deg(\mathrm{P}) \leqslant \deg(\mathrm{Q}) \leqslant 4$.

Exercice 5

Soit f une fonction de classe $\mathcal{C}^2$ sur l'intervalle compact $[a, b]$.

1. Montrer que

$$\int_a^b f(t)\,\mathrm{d}t = \frac{b-a}{2}(f(a)+f(b)) + \frac{1}{2}\int_a^b f''(x)(a-x)(b-x)\,\mathrm{d}x.$$

2. En déduire un encadrement de l'intégrale de f entre a et b en fonction des bornes inférieures et supérieures de f'' sur $[a, b]$.

Exercice 6

Pour $n \in \mathbf{N}$, posons

$$\mathrm{I}_n = \int_0^1 (1-t^2)^n\,\mathrm{d}t.$$

1. Établir une relation de récurrence vérifiée par la suite (I_n).
2. En déduire la valeur de I_n pour tout n.
3. En déduire la valeur de

$$\sum_{k=0}^{n} \frac{(-1)^k}{2k+1}\mathrm{C}_n^k.$$

Exercice 7

Calculer les primitives des fonctions suivantes :

$$\frac{1}{x^2+5},\quad \frac{1}{\sqrt{x^2-5}},\quad \mathrm{e}^x\sin(\mathrm{e}^x),\quad \tan^3 x,\quad \frac{1}{\tan^3 x},$$

$$\frac{2x+3}{(x^2+3x+7)^m},\ m\in\mathbf{N},\quad \frac{\operatorname{Log} x}{x},\quad \frac{\operatorname{ch} x}{\operatorname{sh}^5 x}.$$

Exercice 8

Soit $f : [a, b] \longrightarrow \mathbf{R}$ une fonction de classe $\mathcal{C}^1$, strictement croissante. On définit les deux intégrales I_1 et I_2 par

$$\mathrm{I}_1 = \int_a^b f(t)\,\mathrm{d}t,\quad \mathrm{I}_2 = \int_{f(a)}^{f(b)} f^{-1}(t)\,\mathrm{d}t.$$

1. Justifier l'existence de l'application réciproque f^{-1}. Quelles propriétés possède-t-elle ?
2. Calculer I_2 en fonction de I_1.
3. Faire un dessin faisant apparaître les graphes de f et f^{-1}, et interpréter le résultat ci-dessus graphiquement.

Exercice 9

Déterminer les primitives des fonctions suivantes :

$$\frac{x^4+1}{x(x-1)^3},\quad \frac{1}{(x^4+1)^2},\quad \frac{x}{x^4+x^2+1},\quad \frac{1}{(x-1)(x^2-2x-2)^2},$$

$$\frac{1}{(x+2)(x^2+2x+5)},\quad \frac{2x}{(1-x+x^2)^2},\quad \frac{x^2}{(x-1)^2(x^2+4)}.$$

Exercice 10

Rappelons que l'on définit la suite des intégrales de Wallis par la formule

$$\mathrm{I}_n = \int_0^{\pi/2} \sin^n t \,\mathrm{d}t,\ \ n \geqslant 0.$$

1. Établir une relation de récurrence entre I_n et I_{n-2}. En déduire que si $n = 2p$ est pair alors
$$\mathrm{I}_{2p} = \frac{\pi}{2}\frac{(2p)!}{(2^p p!)^2}$$
et que si $n = 2p+1$ est impair alors
$$\mathrm{I}_{2p+1} = \frac{(2^p p!)^2}{(2p+1)!}.$$
2. Démontrer que la suite (I_n) est décroissante et que $\mathrm{I}_n/\mathrm{I}_{n+1}$ tend vers 1 quand n tend vers l'infini.
3. On définit la suite (u_n) par
$$u_n = (n+1)\mathrm{I}_n\mathrm{I}_{n+1},\ n \geqslant 0.$$
Démontrer que la suite (u_n) est constante et déterminer sa valeur.
4. Démontrer que
$$\mathrm{I}_n \approx \sqrt{\frac{\pi}{2n}}.$$
5. On définit les deux suites (v_n) et (w_n) par
$$v_n = \frac{n^{n+1/2}\mathrm{e}^{-n}}{n!} \quad \text{et} \quad w_n = \mathrm{Log}\left(\frac{v_{n+1}}{v_n}\right).$$
Montrer que
$$w_n = \frac{1}{12n^2} + o(n^{-2}).$$
6. Démontrer que la suite $(\mathrm{Log}(v_n))$ est convergente.

7. Déduire de ce qui précède qu'il existe une constante $\mathrm{C} > 0$ telle que

$$n! \approx \mathrm{C} n^{n+1/2} \mathrm{e}^{-n}.$$

8. En utilisant l'équivalent de la suite (I_n) trouvé plus haut, démontrer que $\mathrm{C} = \sqrt{2\pi}$ et en déduire la *formule de Stirling* :

$$n! \approx \sqrt{2\pi n} \left(\frac{n}{\mathrm{e}}\right)^n.$$

Exercice 11

Soit $f : [a, b] \longrightarrow \mathbf{R}$ une fonction de classe $\mathcal{C}^n$ telle que $f^{(n)}$ soit de plus dérivable sur $[a, b]$.

1. Montrer que la fonction $(b-t)^n f^{(n+1)}(t)$ est intégrable sur $[a, b]$.
2. Montrer, par récurrence sur l'entier n, que la formule suivante, dite *formule de Taylor avec reste intégral*, est vraie :

$$f(b) = f(a) + f'(a)(b-a) + \cdots + f^{(n)}(a)\frac{(b-a)^n}{n!} + \int_a^b f^{(n+1)}(t)\frac{(b-t)^n}{n!}\,\mathrm{d}t.$$

3.(*) Montrer que la formule de la moyenne

$$\exists c \in [a, b], \quad \int_a^b h(x)g(x)\,\mathrm{d}x = g(c)\int_a^b h(x)\,\mathrm{d}x$$

est encore vraie quand on suppose que h est intégrable sur $[a, b]$ et que g satisfait le théorème des valeurs intermédiaires sur tout sous-intervalle $[\alpha, \beta] \subset [a, b]$.

4.(*) Montrer que si g est dérivable sur $[a, b]$ alors g' satisfait le théorème des valeurs intermédiaires sur $[a, b]$.

5. Déduire de ce qui précède une forme du reste dans la formule de Taylor qui ne fasse pas intervenir une intégrale. Le résultat obtenu donne la *formule de Taylor-Lagrange*.

Exercice 12

Calculez, lorsqu'elles sont convergentes, les intégrales suivantes :

$$\int_0^3 \frac{\mathrm{d}x}{\sqrt{9-x^2}}, \quad \int_0^2 \frac{\mathrm{d}x}{2-x}, \quad \int_0^4 \frac{\mathrm{d}x}{(x-1)^2}, \quad \int_0^4 \frac{\mathrm{d}x}{\sqrt[3]{x-1}}.$$

Exercice 13

Calculez, lorsqu'elles sont convergentes, les intégrales suivantes :

$$\int_0^{+\infty} \frac{\mathrm{d}x}{x^2+4}, \quad \int_{-\infty}^2 \mathrm{e}^{2x}\,\mathrm{d}x, \quad \int_1^{+\infty} \sqrt{x}\,\mathrm{d}x, \quad \int_{-\infty}^{+\infty} \frac{\mathrm{d}x}{\mathrm{e}^x + \mathrm{e}^{-x}}, \quad \int_a^{+\infty} \mathrm{e}^{-x}\sin x\,\mathrm{d}x.$$

Exercice 14

Montrer que l'intégrale

$$\mathrm{I} = \int_0^{\pi} \operatorname{Log}(\sin u)\,\mathrm{d}u$$

est convergente. Établir les relations

$$\mathrm{I} = \int_0^{\pi} \operatorname{Log}|\cos u|\,\mathrm{d}u = \int_0^{\pi} \operatorname{Log}|\sin 2u|\,\mathrm{d}u,$$

et en déduire que $\mathrm{I} = -\pi \operatorname{Log} 2$.

Exercice 15

Discuter, suivant les valeurs des nombres réels α, β, la convergence ou la divergence des intégrales suivantes :

$$\int_2^{+\infty} \frac{\operatorname{Log}^{\alpha} t}{t^{\beta}}\,\mathrm{d}t, \quad \int_0^{+\infty} \frac{\operatorname{Log}(1+t^{\alpha})}{t^{\beta}}\,\mathrm{d}t, \quad \int_0^{+\infty} \frac{\sin x}{x^{\alpha}(1+x^{\beta})}\,\mathrm{d}x.$$

Exercice 16 (*)

Soit f une fonction numérique définie sur $[a, +\infty[$. On suppose f positive, décroissante, continue et tendant vers 0 en $+\infty$. Soit par ailleurs g une fonction numérique localement intégrable sur $[a, +\infty[$, telle que l'intégrale indéfinie

$$\psi_g(x) = \int_a^{x} g(x)\,\mathrm{d}x$$

soit bornée indépendamment de x. Montrer qu'alors l'intégrale

$$\int_a^{+\infty} f(x)g(x)\,\mathrm{d}x$$

est convergente (on pourra penser à appliquer la formule de la moyenne pour les intégrales).

12.4. Chapitres 7 et 8

Exercice 1

Déterminer les limites suivantes, en appliquant le théorème de convergence monotone ou le théorème de convergence dominée :

$$\lim_{n\to+\infty} \int_0^{\pi} \frac{n \sin x}{x+n}\,\mathrm{d}x, \quad \lim_{a\to+\infty} \int_0^{\pi} \mathrm{e}^{-a\sin x}\,\mathrm{d}x,$$

$$\lim_{a\to 0} \int_0^{\pi} \mathrm{e}^{-a\sin x}\,\mathrm{d}x, \quad \lim_{n\to+\infty} \int_0^{\pi} \frac{\sin(nx)}{n \sin x}\,\mathrm{d}x$$

(dans le dernier cas, se ramener à une intégration sur l'intervalle $\left[0, \frac{\pi}{2}\right]$).

Exercice 2

Montrer que

$$\mathrm{F}(\alpha) = \int_0^{+\infty} \frac{\mathrm{d}t}{1+t^\alpha}$$

est une fonction continue de la variable $\alpha > 1$. En déduire, après avoir calculé $\mathrm{F}(\alpha)$ pour les α rationnels, que

$$\mathrm{F}(\alpha) = \frac{\pi}{\alpha \sin(\pi/\alpha)}.$$

Exercice 3

Montrer que la fonction

$$\Gamma(x) = \int_0^{+\infty} \mathrm{e}^{-t} t^{x-1}\,\mathrm{d}t$$

est définie pour tout $x > 0$. Montrer que

$$\Gamma(x+1) = x\Gamma(x),$$

et en déduire les valeurs de Γ aux point entiers. Que dire des propriétés de dérivabilité de Γ ?

Exercice 4

Montrer que la fonction

$$f(x) = \int_0^{+\infty} \frac{\arctan(tx)}{t(1+t^2)}\,\mathrm{d}t$$

est définie et dérivable sur $\mathbf{R}$. Calculer f' et en déduire f.

Exercice 5

Soit $f : [0,1] \longrightarrow \mathbf{R}$ une fonction intégrable, on veut déterminer

$$\lim_{n\to+\infty} \int_0^1 n x^n f(x)\,\mathrm{d}x$$

dans les deux cas suivants :

1. On suppose que f' existe et est absolument intégrable ; utiliser une intégration par parties.
2. On suppose que f est absolument intégrable et continue au point 1 ; utiliser la définition de la continuité.

Exercice 6

Soit $f : [1,\mathrm{e}] \longrightarrow \mathbf{R}$ une fonction continue, démontrer que

$$\lim_{n\to+\infty} n \int_1^{1+\frac{1}{n}} f(x^n)\,\mathrm{d}x = \int_1^{\mathrm{e}} \frac{f(x)}{x}\,\mathrm{d}x.$$

Exercice 7

Montrer que pour tout $x \in [0,1[$ on a

$$\int_0^x \operatorname{Log} \frac{1}{1-t} \, \mathrm{d}t = \sum_{n=1}^{+\infty} \frac{x^{n+1}}{n(n+1)},$$

et que la relation est encore vraie pour $x = 1$.

Exercice 8

Soit f_n la fonction définie par $f(x) = x^{2n}\sqrt{\operatorname{ch}(x/n)}$ sur $[-1,1]$. Montrer que la suite (f_n) converge simplement, mais pas uniformément, sur $[-1,1]$, vers un fonction limite que f l'on déterminera. Montrer que

$$\lim_{n \to +\infty} \int_{-1}^{1} f_n(t) \, \mathrm{d}t = 0.$$

Exercice 9

On définit une fonction $f_n : [0,1] \longrightarrow \mathbf{R}$ en posant

$$f_n(x) = \begin{cases} n^2 x & \text{si} \quad 0 \le x < \dfrac{1}{n} \\ 2n - n^2 x & \text{si} \quad \frac{1}{n} \le x < \dfrac{2}{n} \\ 0 & \text{si} \quad \dfrac{2}{n} \le x \le 1 \end{cases}$$

Démontrer que la limite de f_n est nulle mais que la limite de $\int_0^1 f_n(x) \, dx$ n'est pas nulle. En déduire :

1. Qu'il n'existe pas de fonction intégrable $g : [0,1] \longrightarrow \mathbf{R}$ telle que $|f_n| \le g$ pour tout n ;
2. Que la fonction $\sup_{n \in \mathbf{N}^*} f_n$ n'est pas intégrable.

Exercice 10

On pose $\mathrm{I} = [0,1]$ et on y définit la suite de fonctions suivante :

$$f_k(x) = \mathbf{1}_{[1/k, 2/k]}.$$

1. Montrer que f_k converge simplement vers 0 sur I, mais pas uniformément.
2. Montrer que

$$\lim_{k \to +\infty} \int_{-1}^{1} f_k(t) \, \mathrm{d}t = 0.$$

3. Pour tout k, on pose $g_k = kf_k$. Montrer que g_k converge simplement vers 0, mais pas uniformément.
4. Calculer la limite des intégrales de g_k sur I.
5. Montrer que la suite (g_k) ne vérifie pas les hypothèses du théorème de convergence dominée. Que dire du lemme de Fatou ?

Exercice 11

Si E est un ensemble et A une partie de E, rappelons que l'on note $\mathbf{1}_\mathrm{A}$ la fonction indicatrice de A, qui vaut 1 sur A et 0 partout ailleurs.

1. Exprimer, en fonction de $\mathbf{1}_\mathrm{A}$ et $\mathbf{1}_\mathrm{B}$, les fonctions indicatrices des parties suivantes :

$$\mathrm{A} \cap \mathrm{B},\ \ \mathrm{A} \cup \mathrm{B},\ \ \mathrm{A} \bigtriangleup \mathrm{B},\ \ \mathrm{E} - \mathrm{A}.$$

2. Si J est un intervalle borné, exprimer la longueur de J en fonction de $\mathbf{1}_\mathrm{J}$. On appelle cette longueur la *mesure* (de Lebesgue) de J.
3. En déduire qu'on peut définir la mesure d'une union dénombrable d'intervalles ouverts bornés disjoints (penser aux théorèmes de convergence vus en cours).
4. Montrer qu'un ouvert borné est union dénombrable d'intervalles ouverts disjoints. En déduire qu'on peut définir la mesure des ouverts et des fermés bornés de $\mathbf{R}$. Exprimer ces mesures comme des intégrales.
5. Montrer directement en utilisant la définition d'ouvert, que la fonction indicatrice d'un ouvert borné de $\mathbf{R}$ est Lebesgue-intégrable (construire une suite croissante convergeant vers la fonction indicatrice).
6. En examinant les théorèmes de convergence vus en cours, à quelle classe de parties de $\mathbf{R}$ pensez-vous que l'on puisse étendre la mesure ?

Exercice 12 (*)

Rappelons que l'on note $|\mathrm{I}|$ la longueur d'un intervalle de $\mathbf{R}$, et nous généralisons cette notation à une union finie d'intervalles bornés : si $\mathrm{E} = \mathrm{I}_1 \cup \cdots \cup \mathrm{I}_k$, alors $|\mathrm{E}|$ est par définition l'intégrale de la fonction indicatrice de E. Si les I_j sont deux à deux disjoints on a bien sûr $|\mathrm{E}| = |\mathrm{I}_1| + \cdots + |\mathrm{I}_k|$ mais en général on n'a que l'inégalité $|\mathrm{E}| \leqslant |\mathrm{I}_1| + \cdots + |\mathrm{I}_k|$ puisqu'il peut y avoir des recouvrements entre les intervalles I_j. Le lemme d'Austin s'énonce ainsi :

Lemme 12.1 — *Soit $\mathcal{I} = \{\mathrm{I}_1, \ldots, \mathrm{I}_n\}$ une famille finie d'intervalles bornés. Alors il existe une sous-famille $\mathcal{J}$ de $\mathcal{I}$ constituée d'intervalles disjoints deux à deux et telle que*

$$\left| \bigcup_{\mathrm{I} \in \mathcal{J}} \mathrm{I} \right| \geqslant \frac{1}{3} \left| \bigcup_{\mathrm{I} \in \mathcal{I}} \mathrm{I} \right|.$$

Démontrer le lemme d'Austin.

Problème : la fonction de Weierstraß [1]

La première fonction continue partout mais dérivable nulle part a été construite par Weierstraß. Cette fonction a aidé à clarifier les notions de continuité et de dérivabilité, et a obligé les mathématiciens a en donner des définitions précises : auparavant, ceux-ci se contentaient de définitions intuitives, et pensaient qu'une fonction continue était toujours dérivable sauf éventuellement en quelques points. La construction de Weierstraß est venue contredire cette idée intuitive.

Idée de la construction. L'idée de Weierstraß est de partir d'une fonction f_0 qui est parfaitement dérivable, puisqu'il s'agit de la fonction sinus. Puis on perturbe cette

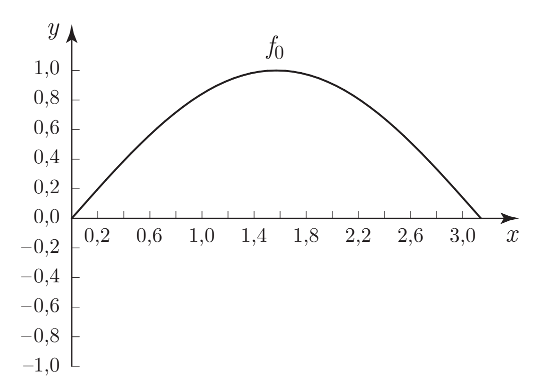


Figure 12.1 – Idée de la construction de la fonction de Weierstraß (1)

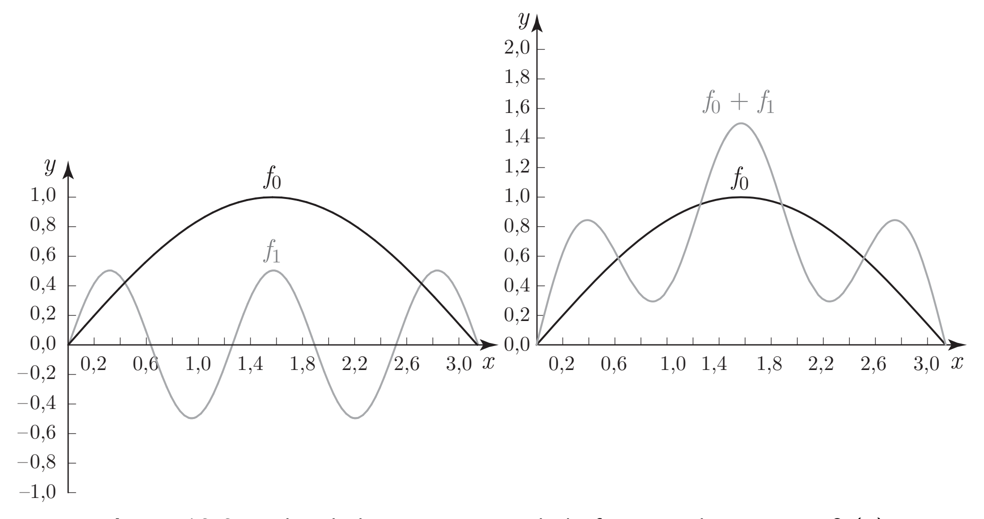


Figure 12.2 – Idée de la construction de la fonction de Weierstraß (2)

1. Exercice et figures créés par Frédéric Le Roux et disponibles sur le site internet *exemaalt*.

première fonction en lui ajoutant une autre sinusoïde, de période 5 fois plus petite, et d'amplitude 2 fois moins grande ; cette deuxième fonction, notée f_1, est représentée sur le premier dessin de la figure 12.2.

Quand on ajoute f_1 à f_0, on obtient une courbe qui zigzague autour du graphe de f_0 (voir le deuxième dessin ; en effet, f_1 est alternativement positive et négative ; quand f_1 est positive, le graphe de $f_0 + f_1$ est situé au-dessus de celui de f_0, et il est situé au-dessous lorsque f_1 est négative ; quand f_1 s'annule, les deux graphes se rencontrent).

Et on recommence : on prend la fonction f_2 qui est une sinusoïde de période 5 fois plus petite et d'amplitude 2 fois moins grande que f_1, et on l'ajoute encore à la fonction $f_0 + f_1$ (voir la figure 12.3). On recommence le processus à l'infini (figure 12.4).

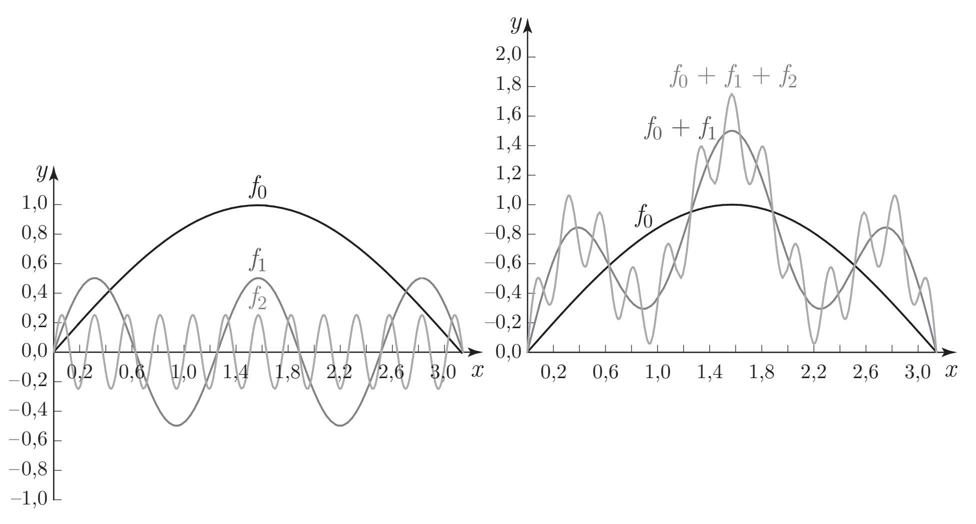


Figure 12.3 – Idée de la construction de la fonction de Weierstraß (3)

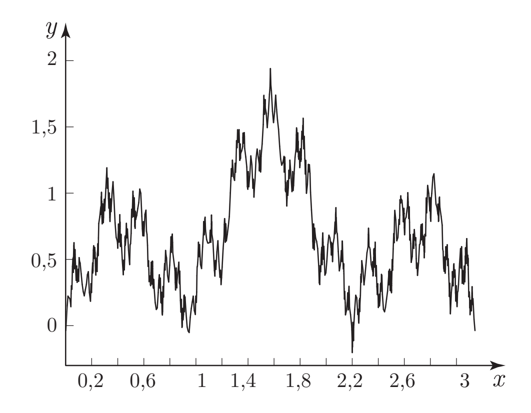


Figure 12.4 – Idée de la construction de la fonction de Weierstraß (∞)

1. Définir précisément f_0, f_1, f_2, …, et F, la fonction limite.
2. Montrer que F est continue.
3. Quelle serait la procédure standard pour montrer que F est dérivable ? Essayez ! Qu'est-ce qui ne marche pas ? Pourquoi ne peut-on pas en conclure que F n'est pas dérivable ?

En fait, la fonction F n'est pas dérivable, mais c'est assez difficile à montrer : on va plutôt examiner une variante de cette première construction, pour laquelle la preuve sera plus facile.

Une version géométrique de la fonction de Weierstraß. Cette fois-ci, on part de la fonction f_0 représentée à la figure 12.5, dont le graphe a la forme d'une dent de hauteur 1/2 et qui est périodique de période 1. La fonction f_1 est alors similaire à f_0, avec huit fois plus de « dents », chaque dent étant deux fois moins élevée que la dent de f_0. Et on recommence : la fonction f_2 a huit fois plus de dents que f_1, de hauteur encore deux fois moins grande, etc.

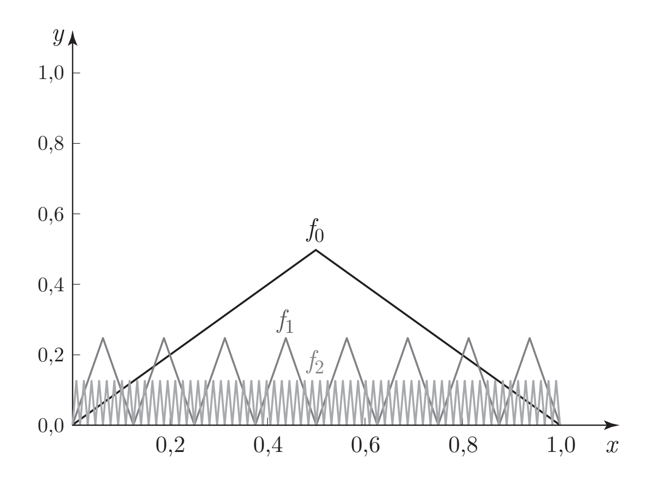


Figure 12.5 – Les nouvelles fonctions f_0, f_1 et f_2

1. Dessiner (le plus précisément possible) $f_0 + f_1$, et (vaguement) $f_0 + f_1 + f_2$. Définir précisément et *en fonction de* f_0, les fonctions $f_1, f_2, \ldots$.
2. Soit F la fonction somme de la série des f_n. Montrer que F est bien définie.
3. Montrer que F est continue.
4. Soient $S_0 = f_0$, $S_1 = f_0 + f_1$, $S_2 = f_0 + f_1 + f_2$, … les sommes partielles de la fonction F.

   Définition — *On dit qu'une fonction est* affine par morceaux *sur l'intervalle* $[0, 1]$ *si cet intervalle est découpé en sous-intervalles, sur chacun desquels la fonction est affine (c.-à-d. que son graphe est un segment).*

Dans ce cas, la dérivée est bien sûr constante sur chaque morceau, et est égale à la pente du segment.
Soit $n > 0$. La fonction S_n est affine par morceaux. Donner les valeurs de la plus grande pente et de la plus petite pente (en valeur absolue). Montrer la propriété **P1** :

il existe une suite $(C_n)_{n>0}$, tendant vers $+\infty$, telle que (pour tout $n > 0$), les valeurs absolues des pentes de tous les morceaux de la fonction S_n sont supérieures ou égales à C_n.

5. Soit n un entier. Trouver tous les endroits où les fonctions F et S_n sont égales ; autrement dit, montrer la propriété **P2** :

   soit $x \in [0,1]$; alors

$$F(x) = S_n(x) \Leftrightarrow x \in \{\dots\} \ : \text{ remplacer les "\dots".}$$

6. Montrer que F n'est pas dérivable en 0.
7. À partir de maintenant, on se donne un nombre $x_0 \in]0,1[$. Montrer la propriété **P3** :

   pour tout entier $n > 0$, il existe deux réels y_n et y'_n tels que :
   (a) y_n et y'_n sont de part et d'autre de x_0 (c.-à-d. que $x_0 \in [y_n, y'_n]$) ;
   (b) la longueur de l'intervalle $[y_n, y'_n]$ est inférieure à $1/8^n$;
   (c) le taux d'accroissement de F entre y_n et y'_n est supérieur (en valeur absolue) à C_n (où (C_n) est la suite définie dans la propriété **P1**).
8. On reprend les notations de la question précédente. Soit, pour chaque entier $n > 0$, t_n le taux d'accroissement, pour F, entre x_0 et y_n et t'_n celui entre x_0 et y'_n. Montrer que l'un des deux taux d'accroissement t_n et t'_n est supérieur (en valeur absolue) à C_n.
9. Montrer que F n'est pas dérivable en x_0.
10. Conclure.

Problème (*) : les fonctions à variation bornée

Soient $I = [a, b]$, φ une fonction réelle définie sur I, et $\mathcal{P} = \{x_0 = a < x_1 < \cdots < x_{n-1} < x_n = n\}$ une partition de I. On note

$$\mathrm{Var}(\varphi, I, \mathcal{P}) = \sum_{k=1}^{n} |\varphi(x_k) - \varphi(x_{k-1})|,$$

et on appelle *variation totale* de φ sur I le nombre (éventuellement infini) suivant :

$$\mathrm{Var}(\varphi, I) = \sup\{\mathrm{Var}(\varphi, I, \mathcal{P}), \mathcal{P} \text{ partition de I}\}.$$

On dit alors que φ est à *variation bornée* si $\mathrm{Var}(\varphi, I)$ est finie.

1. Montrer qu'une fonction monotone est à variation bornée.
2. Soit f la fonction définie sur $[0,1]$ par $f(1) = 0$, et $f(x) = (-1)^{k+1}$ pour tout $x \in [1 - 2^{-(k-1)}, 1 - 2^{-k}[$. Tracer le graphe de f et montrer qu'elle n'est pas à variation bornée.
3. Soit φ une fonction à variation bornée sur $[a, b]$.
   (a) Montrer que $\mathrm{Var}(\varphi, \mathrm{I}) = 0$ si, et seulement si, $\varphi(x) = \varphi(a)$ pour tout x.
   (b) Montrer que $|\varphi(a) - \varphi(b)| \leqslant \mathrm{Var}(\varphi, \mathrm{I})$.
   (c) En déduire qu'une fonction à variation bornée est bornée.
4. Montrer qu'une fonction Lipschitzienne est à variation bornée, mais que la réciproque est fausse.
5. (a) Montrer que $\mathrm{Var}(\sin, [0, \pi]) = 2$.
   (b) Soit $\varphi(x) = x \sin(\pi x^{-1})$ sur $]0, 1]$ et $\varphi(0) = 0$. Montrer que φ n'est pas à variation bornée.
   (c) Évaluer $\mathrm{Var}(|\cos|, [0, 6\pi])$.
6. Soit $\mathbf{BV}(\mathrm{I})$ l'ensemble des fonctions à variation bornée sur I. Montrer que $\mathbf{BV}(\mathrm{I})$ est une $\mathbf{R}$-algèbre pour la somme et le produit usuels des fonctions.
7. Soit $c \in \,]a, b[$. Montrer que $\varphi \in \mathbf{BV}(\mathrm{I})$ si, et seulement si, $\varphi \in \mathbf{BV}([a, c]) \cap \mathbf{BV}([c, b])$. Montrer qu'alors
$$\mathrm{Var}(\varphi, [a, b]) = \mathrm{Var}(\varphi, [a, c]) + \mathrm{Var}(\varphi, [c, b]).$$
8. On va ici montrer un important théorème dû à Camille Jordan, qui caractérise les fonctions à variation bornée comme étant les fonctions différence de deux fonctions croissantes.
   (a) Montrer que si φ est la différence de deux fonctions monotones, elle est à variation bornée.
   (b) Supposons maintenant $\varphi \in \mathbf{BV}(\mathrm{I})$. Montrer que les fonctions V et W définies ci-dessous sont bien définies et croissantes :
$$\mathrm{V}(x) = \mathrm{Var}(\varphi, [a, x]), \ \mathrm{W}(x) = \mathrm{V}(x) - \varphi(x).$$
   En déduire le théorème de Jordan.
9. Soit f une fonction absolument intégrable sur $[a, b]$, c'est-à-dire $f \in \mathcal{L}(\mathrm{I})$. Montrer que son intégrale indéfinie
$$\mathrm{F}(x) = \int_a^x f(t) \, \mathrm{d}t$$
est à variation bornée. La réciproque est vraie, mais un peu délicate à démontrer (essayez !).

Problème : autour de la moyenne arithmético-géométrique

Pour deux nombres réels a, b satisfaisant à $0 < b \leqslant a$ on définit deux suites (a_n) et (b_n) de la manière suivante :
$$a_0 = a, \ b_0 = b, \ a_{n+1} = \frac{a_n + b_n}{2}, \ b_{n+1} = \sqrt{a_n b_n}.$$

1. Montrer que pour tous $a, b \in \mathbf{R}_+$ on a $\sqrt{ab} \leqslant (a+b)/2$.
2. Montrer que pour tout n on a $b_n \leqslant a_n$, que la suite (a_n) est décroissante et la suite (b_n) croissante.
3. Montrer que pour tout n on a
$$a_{n+1} - b_{n+1} \leqslant \frac{a_n - b_n}{2}$$
et en déduire que les suites (a_n) et (b_n) convergent vers une même limite. Nous la noterons $\mathrm{M}(a, b)$; elle s'appelle la moyenne arithmético-géométrique de a et b. C'est une notion due à Gauß qui apparaît comme par enchantement dans de nombreux problèmes mathématiques.
4. On pose $c_n = \sqrt{a_n^2 - b_n^2}$; montrer que l'on a pour tout n
$$c_{n+1} \leqslant \frac{c_n^2}{4\mathrm{M}(a, b)}.$$
À quelles vitesses les suites (a_n) et (b_n) convergent-elles ?
5. Soit f l'application définie sur l'intervalle $[0, 1]$ par
$$f(t) = \frac{2at}{a + b + (a-b)t^2}.$$
Montrer que f réalise un difféomorphisme de classe $\mathcal{C}^1$ de $[0, 1]$ sur lui-même.
6. Montrer qu'en posant
$$\sin\varphi = \frac{2a \sin\varphi'}{a + b + (a-b)\sin^2\varphi'}$$
on définit correctement un difféomorphisme de classe $\mathcal{C}^1$ de $[0, \pi/2]$ sur lui-même. Montrer qu'alors on a les formules
$$\cos\varphi = \frac{2\cos\varphi'(a_1^2\cos^2\varphi' + b_1^2\sin^2\varphi')}{a + b + (a-b)\sin^2\varphi'}$$
et
$$\frac{1}{\sqrt{a^2\cos^2\varphi + b^2\sin^2\varphi}} = a\frac{a + b - (a-b)\sin^2\varphi'}{a + b + (a-b)\sin^2\varphi'}.$$
7. On pose
$$\mathrm{I}(a, b) = \int_0^{\pi/2} \frac{\mathrm{d}\varphi}{\sqrt{a^2\cos^2\varphi + b^2\sin^2\varphi}}.$$
Montrer que $\mathrm{I}(a, b) = \mathrm{I}(a_1, b_1)$.

8. En déduire l'identité
$$\mathrm{M}(a,b)\mathrm{I}(a,b) = \frac{\pi}{2}.$$

9. Nous notons ici $\mathcal{L}$ la lemniscate, courbe plane donnée en coordonnées polaires par $r^2 = \cos 2\theta$. Étudier la courbe $\mathcal{L}$ et la tracer. Montrer qu'elle est rectifiable ; nous noterons l sa longueur. Montrer que
$$l = 4\int_0^{\pi/4} \frac{\mathrm{d}\theta}{\sqrt{\cos 2\theta}} = 4\int_0^{\pi/2} \frac{\mathrm{d}\varphi}{\sqrt{2\cos^2\varphi + \sin^2\varphi}}$$
et en déduire que
$$l = \frac{2\pi}{\mathrm{M}(\sqrt{2},1)}.$$

12.5. Chapitre 9

Exercice 1

Soit $\alpha \in \mathbf{R}$. Déterminer les coefficients de Fourier de la fonction 2π-périodique f définie par $f(t) = \cos(t+\alpha)$.

Exercice 2

Déterminer les coefficients de Fourier de la fonction 2π-périodique égale à 1 sur $[0,\pi[$ et -1 sur $[\pi, 2\pi[$. En quels points de $\mathbf{R}$ la série converge-t-elle simplement ?

Exercice 3

Soit α un nombre réel. Considérons les deux fonctions 2π-périodiques s_α et c_α définies sur $[-\pi,\pi[$ par $s_\alpha(x) = \sin(\alpha x)$ et $c_\alpha(x) = \cos(\alpha x)$. Calculer les coefficients de Fourier de s_α et c_α. En déduire que pour tout $x \notin \mathbf{Z}$ on a
$$\pi \operatorname{cotan}(\pi x) = \frac{1}{x} + 2x\sum_{n=1}^{\infty} \frac{1}{x^2-n^2}$$
et
$$\frac{\pi}{\sin \pi x} = \frac{1}{x} + 2x\sum_{n=1}^{\infty} \frac{(-1)^n}{x^2-n^2}.$$

Exercice 4

La fonction $x \mapsto \operatorname{Log}|\sin(x/2)|$ est-elle 2π-périodique ? Si oui, déterminer ses coefficients de Fourier.

Exercice 5 (*)

Soit f une fonction 2π-périodique et Lebesgue-intégrable sur $[0, 2\pi]$. Nous noterons $\mathrm{S}_n(x)$ la somme de la série de Fourier réelle de f au rang n :

$$\mathrm{S}_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx),$$

où a_k, b_k sont les coefficients de Fourier réels de f. On définit la moyenne de Cesaro de la suite $\mathrm{S}_n(x)$ par

$$\sigma_n(x) = \frac{1}{n}\sum_{k=0}^{n} \mathrm{S}_n(x).$$

1. Montrer que

$$\sigma_n(x) = \frac{1}{2\pi n}\int_0^{\pi}[f(x+u) + f(x-u)]\frac{\sin^2(nu/2)}{\sin^2(u/2)}\,\mathrm{d}u.$$

2. Montrer que

$$\int_0^{\pi} \frac{\sin^2(nu/2)}{\sin^2(u/2)}\,\mathrm{d}u = n\pi.$$

3. En vous inspirant des méthodes utilisées pour démontrer le critère de Dini, démontrer que $\sigma_n(x)$ converge vers $f(x)$ en tout point x où f est continue.
4. En déduire que si f est continue et que tous ses coefficients de Fourier sont nuls, alors $f = 0$.

Exercice 6

On considère une corde vibrante dont les extrémités sont situées en 0 et en ℓ. On pince la corde de manière à ce qu'elle ait, au temps $t = 0$, la forme donnée par la fonction suivante : $u(x, 0) = \alpha x$ pour $0 \leqslant x \leqslant d$ et $u(x, 0) = \alpha dx/(d - \ell) + \alpha d\ell/(\ell - d)$ pour $d \leqslant x \leqslant \ell$. On lâche alors la corde de sorte que sa vitesse initiale soit nulle en tout point. Déterminer l'évolution ultérieure de la corde sachant qu'elle obéit à l'équation des cordes vibrantes.

12.6. Chapitre 10

Exercice 1

Calculer les intégrales multiples suivantes :

$$\int_{\mathrm{T}} xy\,\mathrm{d}x\,\mathrm{d}y, \quad \mathrm{T} = \left\{(x, y) \in (0, 1)^2, y \leqslant x\right\};$$

$$\int_{\mathrm{D}} \frac{xy}{1 + x^2 + y^2}\,\mathrm{d}x\,\mathrm{d}y, \quad \mathrm{D} = \left\{(x, y) \in [0, 1]^2, x^2 + y^2 \geqslant 1\right\};$$

$$\int_{\mathrm{D}'} \frac{1}{y\cos x + 1}\,\mathrm{d}x\,\mathrm{d}y, \quad \mathrm{D}' = \left[0, \frac{\pi}{2}\right] \times \left[0, \frac{1}{2}\right].$$

Exercice 2

Soit (x,y) un point de $\mathbf{R}^2$ différent de l'origine.

1. Rappeler ce que sont les coordonnées polaires (r,θ) de (x,y). On note $(t,\theta) = \mathrm{P}(x,y)$
2. Montrer que l'application P définie sur $\mathbf{R}^2$ privé du demi axe des x négatif, et à valeurs dans $[0,+\infty[\times]-\pi,\pi[$ est un difféomorphisme. Calculer son jacobien.
3. Si f est une fonction définie sur $\mathbf{R}^2$ et si $(r,\theta) = (x,y)$, on utilise un abus de notation consistant à écrire $f(x,y) = f(r,\theta)$.
   (a) En quoi le fait que P ne soit définie que sur $\mathbf{R}$ privé d'une demi-droite n'est pas gênant si l'on s'intéresse à l'intégration des fonctions sur $\mathbf{R}^2$?
   (b) Exprimer l'intégrale de f sur $\mathbf{R}^2$ si l'on travaille en coordonnées polaires.
   (c) On dit que f est une fonction radiale si $f(r,\theta)$ ne dépend que de r. Exprimer dans ce cas l'intégrale de f en coordonnées polaires.
   (d) En déduire le moment d'inertie d'une surface plane présentant une symétrie de révolution.

Exercice 3

Soit $\mathrm{D}_r = \{(x,y) \in \mathbf{R}^2, x^2+y^2 \leqslant r\}$ le disque de rayon r, et $\mathrm{K}_r = [-r,r]\times[-r,r]$ le carré de côté $2r$.

1. Calculer
$$\int_{\mathrm{D}_r} \mathrm{e}^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y$$
en passant en coordonnées polaires.
2. Trouver une constante $c > 0$ telle que, pour tout $r > 0$, on ait $\mathrm{D}_r \subset \mathrm{K}_r \subset \mathrm{D}_{cr}$.
3. En déduire la valeur de
$$\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2}\,\mathrm{d}t.$$

On peut démontrer (c'est difficile, on utilise pour cela ce que l'on appelle la théorie de Galois différentielle) que la fonction e^{-t^2} n'admet pas de primitive exprimable par des fonctions usuelles. Mais le passage par des intégrales doubles permet de calculer l'intégrale sur $\mathbf{R}$!

Exercice 4

On se place dans l'espace $\mathbf{R}^3$. Soient $a > \mathrm{R} > 0$ et soit, dans le plan $\{x=0\}$ le disque centré en $(0,a,0)$ de rayon R. On fait alors tourner ce disque autour de l'axe des z pour obtenir le solide de révolution $\mathrm{T}_{a,\mathrm{R}}$ que l'on appelle un tore.

1. Effectuer un dessin de $\mathrm{T}_{a,\mathrm{R}}$.
2. Calculer le volume de $\mathrm{T}_{a,\mathrm{R}}$.
3. On considère les coordonnées cylindriques (r,θ,z) sur $\mathbf{R}^3$ et l'application d : $\mathbf{R}^3-\{0\}$ dans lui-même définie par $d(r,\theta,z) = d(r,2\theta,z)$. Déterminer le jacobien de cette transformation vue comme transformation de $\mathbf{R}^3$ en coordonnés cylindriques.

Montrer que d est une application différentiable de $\mathrm{T}_{a,\mathrm{R}}$ dans lui-même dont le jacobien est partout non nul, mais qui est telle que tout point admet *deux* préimages. Une telle situation est-elle possible sur un intervalle ?

Exercice 5

Calculer le volume de la boule unité $\mathbf{B}_d$ de $\mathbf{R}^d$:

$$\mathbf{B}_d = \big\{(x-1,\dots,x_d) \in \mathbf{R}^d,\ \ x_1^2+\dots+x_d^2 \leqslant 1\big\}.$$

Exercice 6

On considère la partie compacte Δ de $\mathbf{R}^3$ définie par

$$\Delta = \big\{(x_1,x_2,x_3) \in \mathbf{R}_+^3,\ \ x_1+x_2+x_3 \leqslant 1\big\}.$$

1. Dessiner le compact Δ et calculer son volume.
2. Trois entiers positifs q_1, q_2, q_3 étant donnés, calculer l'intégrale

$$\iiint_\Delta x_1^{q_1} x_2^{q_2} x_3^{q_3}\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3.$$

Exercice 7 (*)

On considère une fonction f de classe $\mathcal{C}^2$ définie sur le disque unité fermé $\mathrm{D}_1 = \{(x,y) \in \mathbf{R}^2, x^2+y^2 \leqslant 1$ de $\mathbf{R}^2$. On rappelle que le laplacien de f est défini par

$$\Delta f = \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y}.$$

Pour $0 < r < 1$ on note

$$\mathrm{M}_r(f) = \frac{1}{\pi r^2}\iint_{\mathrm{D}_r} f(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

la moyenne de f sur le disque de rayon r. En utilisant la formule de Taylor, démontrer que

$$\Delta f(0,0) = \lim_{r\to 0}\frac{2}{r^2}(\mathrm{M}_r(f) - f(0,0)).$$

12.7. Chapitre 11

Exercice 1

Soit V un $\mathbf{R}$-espace vectoriel muni d'une semi-norme $\|\cdot\|$. Montrer que l'ensemble des vecteurs de semi-norme nulle est un sous-espace fermé de V.

Exercice 2

Soit $(r_n)_{n\in\mathbf{N}}$ une énumération des nombres rationnels compris entre 0 et 1. Montrer que la fonction définie sur $[0,1]$ par

$$f(x) = \sum_{k=0}^{\infty} 2^{-k} \frac{1}{\sqrt{|x-r_k|}}$$

est finie presque partout. Que peut-on dire de l'intégrale de f^2 sur un intervalle longueur non nulle ?

Exercice 3

Soient $\Omega \subset \mathbf{R}^d$ un ouvert de mesure de Lebesgue *finie* et p, q deux nombres réels tels que $1 \leqslant p \leqslant q$. Montrer que $\mathrm{L}^q(\Omega) \subset \mathrm{L}^p(\Omega)$. Ce résultat est-il vrai sans l'hypothèse de mesure finie ? Démontrer enfin que

$$\sup\{\|f\|_{\mathrm{L}^p}, f \in \mathrm{L}^q(\Omega)\} = |\Omega|^{p^{-1}-q^{-1}}.$$

Exercice 4

Soit $p \geqslant 1$. Nous allons construire une fonction sur $\mathbf{R}^d$ qui est dans $\mathrm{L}^p(\mathbf{R}^d)$ pour *un seul* p.

1. Montrer qu'il existe une fonction f dont le support est inclus dans la boule unité $\mathbf{B}$ de $\mathbf{R}^d$, qui est dans L^p mais dans aucun L^q pour $q > p$.
2. Rappelons qu'on note ℓ^q l'ensemble des suites $(x_n)_{n\in\mathbf{N}}$ de nombres réels telles que

$$\sum_{k=0}^{\infty} |x_k|^q < +\infty.$$

   Démontrer qu'il existe une suite $\mathbf{x} = (x_n)$ qui est ℓ^p mais dans aucun ℓ^q pour $q < p$.
3. Soit $(\mathrm{B}_n)_{n\in\mathbf{N}}$ une famille de boules de $\mathbf{R}^d$, disjointes deux à deux et toutes de rayon 1. Démontrer que la fonction

$$g = f + \sum_{k=0}^{\infty} x_k \mathbf{1}_{\mathrm{B}_k}$$

   est dans L^p mais dans aucun L^q pour $q \neq p$.

Exercice 5

Soit X un ensemble. Une famille de parties $\mathcal{B}$ de X est appelé une tribu si pour $\mathcal{B}$ contient $\emptyset, \mathrm{X}$, est stable par passage au complémentaire et est stable par union dénombrable.

1. Montrer que l'ensemble $\mathbf{M}$ des parties Lebesgue-mesurables de $\mathbf{R}^d$ est une tribu.
2. Montrer que l'intersection d'une famille de tribus est encore une tribu. En déduire la notion de *tribu engendrée par une famille de parties de* X.
3. Soient $A \subset X$ une partie de X et $\mathcal{R} \subset \mathcal{P}(X)$ un famille de parties de X. On note $T(\mathcal{R})$ la tribu engendrée par $\mathcal{R}$ et $\mathcal{R}_A = \{Y \cap A,\ Y \in \mathcal{R}\}$. Démontrer que $T(\mathcal{R}_A) = T(\mathcal{R})_A$.
4. Une famille $\mathcal{A}$ de parties de X est une *algèbre booléenne* si elle contient $\emptyset, X$, est stable par passage au complémentaire et est stable par union *finie.* Une famille $\mathcal{M}$ de parties de X est une *classe monotone* si :
   - pour toute famille dénombrable décroissante $E_1 \supset E_2 \supset \cdots$ d'éléments de $\mathcal{M}$, on a
   $$\bigcap_{i\in\mathbf{N}} E_i \in \mathcal{M}$$
   - et si pour toute famille dénombrable croissante $F_1 \subset F_2 \subset \cdots$ d'éléments de $\mathcal{M}$ on a
   $$\bigcup_{i\in\mathbf{N}} F_i \in \mathcal{M}.$$

   Démontrer le *théorème des classes monotones* : si $\mathcal{A}$ est une algèbre booléenne et si $\mathcal{M}$ est une classe monotone contenant $\mathcal{A}$, alors $\mathcal{M}$ contient la tribu engendrée par $\mathcal{A}$.

Exercice 6

Soient X un ensemble et $\mathcal{B} \subset \mathcal{P}(X)$ une tribu. On appelle *mesure de probabilité* définie sur $(X, \mathcal{B})$ toute application $\mu : \mathcal{B} \longrightarrow \mathbf{R}$ telle que
- pour tout $A \in \mathcal{B}$, $\mu(A) \geqslant 0$;
- pour toute famille dénombrable $(A_i)_{i\in\mathbf{N}}$ d'éléments de $\mathcal{B}$ *deux à deux disjointes* on a
$$\mu\left(\bigcup_{i\in\mathbf{N}} A_i\right) = \sum_{i\in\mathbf{N}} \mu(A_i)\,;$$
- $\mu(X) = 1$.

1. Donner quelques exemples de mesures de probabilités.
2. Montrer les propriétés suivantes :
   - $\mu(A \cup B) \leqslant \mu(A) + \mu(B)$;
   - $\mu(\emptyset) = 0$;
   - si $A \subset B$ alors $\mu(A) \leqslant \mu(B)$;
   - si $(A_i)_{i\in\mathbf{N}}$ est une suite croissante d'éléments de $\mathcal{B}$, alors
   $$\mu\left(\bigcup_{i\in\mathbf{N}} A_i\right) = \lim_{n\to+\infty} \mu(A_i).$$

3. Une famille Π de parties de X est un π-système si $\mathrm{A}, \mathrm{B} \in \Pi$ entraîne $\mathrm{A} \cap \mathrm{B} \in \Pi$. Soient Π un π-système constitué d'éléments de $\mathcal{B}$, μ_1 et μ_2 deux mesures de probabilité telles que μ_1 et μ_2 coïncident sur les éléments de Π. Montrer qu'alors μ_1 et μ_2 coïncident sur la tribu engendrée par Π : c'est le théorème des π-systèmes.
4. On pose $\mathrm{X} = \{0,1\}^{\mathbf{N}}$. La donnée d'un entier $n \geqslant 0$ et de tirages $\alpha_0, \ldots, \alpha_{n-1} \in \{0,1\}$ permet de définir un *cylindre* :

$$\mathrm{C} = \big\{(\beta_i)_{i \in \mathbf{N}} \in \mathrm{X},\ \beta_0 = \alpha_0, \ldots, \beta n - 1 = \alpha_{n-1}\big\}.$$

Montrer que l'ensemble des cylindres forme un π-système. Si C est un cylindre comme ci-dessus, on définit sa mesure par $\mu(\mathrm{C}) = 2^{-n}$. Le théorème des π-systèmes assure que si μ s'étend à la tribu engendrée par les cylindres, alors cette extension est unique. L'existence de l'extension peut s'obtenir grâce au théorème d'extension de Carathéodory ou au théorème de consistence de Kolmogoroff. Nous renvoyons la lectrice à tout bon livre de théorie des probabilités où ces théorèmes sont démontrés.

Exercice 7

Soient $p > 1$ un nombre réel et $\mathrm{U} =]0, +\infty[$. Si $f \in \mathrm{L}^p(\mathrm{U})$ on note $\mathrm{T}(f)$ la fonction définie sur U par

$$\mathrm{T}f(x) = \frac{1}{x} \int_0^x f(t)\,\mathrm{d}t.$$

1. Montrer que la fonction $x \mapsto \mathrm{T}f(x)$ est bien définie pour tout $x > 0$.
2. Montrer que si f est continue à support compact dans U, alors f et $\mathrm{T}f$ sont liées par une équation différentielle que l'on explicitera.
3. On suppose ici f continue à support compact. Montrer que l'on a

$$\|\mathrm{T}(f)\|_{\mathrm{L}^p} \leqslant \frac{p}{p-1} \|f\|_{\mathrm{L}^p}.$$

On pourra commencer par traiter le cas où f est de signe constant.
4. Montrer que si (f_n) converge vers f dans L^p alors $\mathrm{T}(f_n)$ converge simplement vers $\mathrm{T}(f)$.
5. Si $f \in \mathrm{L}^1(\mathrm{U})$, a-t-on $\mathrm{T}(f) \in \mathrm{L}^1(\mathrm{U})$?

Exercice 8

Démontrer que la fonction h définie sur $\mathbf{R}$ par

$$h(x) = \sum_{n \geqslant 0} 2^{-n} \cos(3^n x)$$

est continue et nulle part dérivable.

Exercice 9

Montrer que si $f : \mathbf{R} \longrightarrow \mathbf{R}$ est de classe $\mathcal{C}^k$ et 2π-périodique alors sa série des coefficients de Fourier satisfait à

$$\sum_{n \in \mathbf{Z}} |n|^{2k} |c_n| < +\infty.$$

Les problèmes donnés maintenant sont longs et assez difficiles, mais le lecteur courageux les lira avec profit.

Problème : inégalité maximale de Hardy et Littlewood

Le but de ce problème est de démontrer l'inégalité maximale de Hardy et Littlewood, avatar du lemme maximal utilisé dans la démonstration du théorème ergodique.

1. Une version discrète.
   (a) Soit $a_1, \ldots, a_{\mathrm{N}-1}$ des nombres réels positifs ou nuls. Pour tout $1 \leqslant n \leqslant \mathrm{N}_1$ et tout $0 \leqslant k \leqslant \mathrm{N} - 1 - n$ on note $s_{n,k}$ la moyenne des k termes de la suite (a_i) démarrant en a_n :

   $$s_{n,k} = \frac{1}{k} \sum_{j=0}^{k} a_{n+j}.$$

   On note enfin

   $$a_n^* = \max_k \{s_{n,k}\}.$$

   On se donne un nombre réel $\lambda > 0$ et l'on note $\mathrm{E} = \{n,\ a_n^* \geqslant \lambda\}$ ainsi que $b_n = a_n - \lambda$ et $\mathbf{1}_\mathrm{E}$ la fonction indicatrice de E. Montrer que pout tout $1 \leqslant n \leqslant \mathrm{N} - 1$ on a $b_n \mathbf{1}_\mathrm{E}(n) \geqslant b_n$.
   (b) En vous inspirant le la démonstration du lemme maximal 11.33 utilisé dans la démonstration du théorème ergodique, montrer l'existence d'une fonction $\tau : \{1, \ldots, \mathrm{N} - 1\} \longrightarrow \{1, \ldots, \mathrm{N} - 1\}$ telle que pour tout n on ait

   $$\sum_{j=0}^{\tau(n)-1} b_{n+j} \mathbf{1}_\mathrm{E}(n + j) \geqslant 0.$$

   (c) Toujours en vous inspirant de la preuve du lemme maximal 11.33, montrer que l'on a l'inégalité

   $$\sum_{n=1}^{\mathrm{N}-1} b_n \chi_\mathrm{E}(n) \geqslant 0.$$

   (d) En déduire ce que nous appellerons l'inégalité maximale discrète :

   $$\mathrm{Card}\{n, a_n^* \geqslant \lambda\} \leqslant \frac{1}{\lambda} \sum_{n=1}^{\mathrm{N}-1} a_n.$$

2. Nous allons maintenant passer à la démonstration de l'inégalité de Hardy-Littlewood proprement dite. On se donne une fonction $f \in \mathrm{L}^1([0,1])$ et pour tout $x \in [0,1[$ on pose
$$f^*(x) = \sup\left\{\frac{1}{h}\int_0^1 |f(t)|\,\mathrm{d}t,\ h > 0\right\}.$$
Pour tout $n \geqslant 1$ on pose
$$f_n(x) = n\int_x^{x+n^{-1}} |f(t)|\,\mathrm{d}t.$$
On se donne enfin un nombre réel $\lambda > 0$ et l'on pose $\mathrm{E}_\lambda^* = \{t,\ f^*(t) \geqslant \lambda\}$. Montrer que f^* est une fonction mesurable et en déduire que E_λ^* est une partie mesurable de $[0,1]$. Le but de la fin de ce problème est de démontrer l'inégalité maximale suivante :
$$|\mathrm{E}_\lambda^*| \leqslant \frac{1}{\lambda}\|f\|_{\mathrm{L}^1}.$$
3. Montrer que si f est continue et positive on a $|\mathrm{E}_\lambda^*| \leqslant \lambda^{-1}\|f\|_{\mathrm{L}^1}$. On pourra introduire une primitive de f et penser à l'inégalité de Tchebicheff.
4. On revient au cas général de $f \in \mathrm{L}^1$ mais en supposant encore $f \geqslant 0$. Pour $x \in [0, \mathrm{N}^{-1}]$ et $0 < k < \mathrm{N}$ on pose
$$f_\mathrm{N}(x,k) = \mathrm{N}\int_{x+(k-1)\mathrm{N}^{-1}}^{x+k\mathrm{N}^{-1}} f(t)\,\mathrm{d}t$$
et
$$\psi_\mathrm{N}(x) = \sup\left\{\frac{\mathrm{N}}{n}\int_x^{x+n\mathrm{N}^{-1}} f(t)\,\mathrm{d}t,\ 1 \leqslant n \leqslant \mathrm{N}-1\right\}.$$
Montrer que l'on a l'identité suivante :
$$|\{x,\ \psi_\mathrm{N}(x) \geqslant \lambda\}| = \int_0^{\mathrm{N}^{-1}} \mathrm{Card}\{k,\ f_\mathrm{N}^*(t,k) \geqslant \lambda\}\,\mathrm{d}t.$$
En déduire que $|\{x,\ \psi_\mathrm{N}(x) \geqslant \lambda\}| \leqslant \|f\|_{\mathrm{L}^1}$ puis que $\|x,\ f(x) \geqslant \lambda\}| \leqslant \|f\|_{\mathrm{L}^1}$.
5. Démontrer l'inégalité maximale dans le cas général de $f \in \mathrm{L}^1$.

Problème : transformations expansives de l'intervalle

On va s'intéresser ici à certaines transformations de l'intervalle qui ont des propriétés proches de la transformation de Gauß $\mathrm{G}(x) = \{x^{-1}\}$. On pose $\mathrm{I} = [0,1]$ et l'on se donne une famille dénombrable $(\mathrm{I}_n)_{n\in\mathbf{N}}$ de sous-intervalles compacts de $[0,1]$ dont on suppose qu'elle satisfait aux deux propriétés suivantes :
– les I_n sont d'intérieur non vides et d'intérieurs deux à deux disjoints ;
– l'union S de tous les I_n est de mesure 1.

On se donne enfin une transformation $f : \mathrm{I} \longrightarrow \mathrm{I}$ qui satisfait aux trois conditions suivantes :

– pour tout $j \in \mathbf{N}$, f définit un difféomorphisme de classe $\mathcal{C}^1$ entre I_j et I. On note g_j sont inverse ;
– il existe un nombre $\lambda > 1$ tel que pour tout $x \in \mathrm{S}$ on ait $|f'(x)| > \lambda$;
– il existe $0 < \alpha \leqslant 1$ et $c_0 > 0$ tels que pour tout $j \in \mathbf{N}$ et tout $x, y \in \mathrm{I}_j$ on ait

$$\left|\frac{f'(x)}{f'(y)}\right| \leqslant \mathrm{e}^{c_0|x-y|^\alpha}.$$

La seconde condition signifie que sur chaque intervalle I_j la transformation f est « dilatante » et la seconde condition dit que sur chaque intervalle I_j la transformation f n'est pas trop « distordue ».

1. Pour $\mathrm{J} = (j_1, \ldots, j_n) \in \mathbf{N}^n$ on note $g_\mathrm{J} = g_{j_1} \circ g_{j_n}$ et $\mathrm{I}_j = g_\mathrm{J}(\mathrm{I})$. Montrer que la famille d'intervalles $(\mathrm{I}_\mathrm{J})_{\mathrm{J}\in\mathbf{N}^n}$ satisfait aux mêmes conditions que la famille $(\mathrm{I}_j)_{j\in\mathbf{N}}$.
2. Montrer que pour tous $x, y \in \mathrm{I}$ et tout $\mathrm{J} \in \mathbf{N}^n$ on a

$$\left|\mathrm{Log}\,\frac{|g'_\mathrm{J}(x)|}{|g'_\mathrm{J}(y)|}\right| \leqslant c_1|x-y|^\alpha,$$

   où $c_1 = c_0(\lambda^\alpha - 1)^{-1}$.
3. En déduire que la fonction

$$\mathrm{S}_n(x) = \sum_{\mathrm{J}\in\mathbf{N}^n} |g'_\mathrm{J}(x)|$$

   est bien définie et qu'elle satisfait aux deux conditions suivantes

$$(\star) \qquad \forall x, y \in \mathrm{I},\ \mathrm{e}^{-c_1} \leqslant \mathrm{S}_n(x) \leqslant \mathrm{e}^{c_1} \quad \text{et} \quad |\mathrm{S}_n(x) - \mathrm{S}_n(y)| \leqslant c_2|x-y|^\alpha,$$

   où c_2 est une constante à déterminer. On pourra penser à appliquer l'égalité des accroissements finis à la fonction logarithme sur un intervalle de bornes $\mathrm{Log}\, g'_\mathrm{J}(x)$ et $\mathrm{Log}\, g'_\mathrm{J}(y)$.
4. Montrer que pour tout ensemble mesurable $\mathrm{A} \subset \mathrm{I}$ on a

$$|f^{-n}(\mathrm{A})| = \int_\mathrm{A} \mathrm{S}_n(x)\,\mathrm{d}x.$$

5. On pose

$$\mathrm{H}_n(x) = \frac{1}{n}\sum_{k=0}^{n-1} \mathrm{S}_k(x).$$

   Montrer que H_n satisfait aux conditions $(\star)$. En déduire qu'il existe une sous-suite n_k et une fonction H continue sur I satisfaisant aux conditions $(\star)$ telle que $(\mathrm{H}_{n_k})_k$ converge uniformément vers H sur I.
6. Montrer que f préserve la mesure à densité $\mathrm{H}(x)\,\mathrm{d}x$.

Problème : le théorème ergodique de Birkhoff appliqué aux fractions continues

Le but de ce problème est d'étudier certaines propriétés statistiques élémentaires du développement en fraction continue des nombres réels. Nous devons pour cela introduire quelques notations.

Étant donnée $(a_n)_{n\geqslant 1}$ une suite (finie ou infinie) d'entiers naturels, on définit par récurrence deux autres suites d'entiers (p_n) et (q_n) de la manière suivante :

$$p_{-1} = 1, p_0 = 0, p_n = a_n p_{n-1} + p_{n-2}$$

et

$$q_{-1} = 0, q_1 = 1, q_n = a_n q_{n-1} + q_{n-2}.$$

Si $x \in [0, 1[$ est un nombre réel, on définit

$$\mathrm{T}(x) = \left\{ \frac{1}{x} \right\} \text{ si } x \neq 0,\ \mathrm{T}(0) = 0,$$

où $\{x\}$ désigne la partie fractionnaire $x - [x]$ de x. Ainsi T est une application de $[0, 1[$ dans lui-même : c'est la *transformation de Gauß*. On pose également

$$a(x) = \left[\frac{1}{x} \right] \text{ si } x \neq 0, \quad a(0) = \infty$$

et enfin, pour tout $n = 1, 2 \ldots$

$$a_n(x) = a(\mathrm{T}^{n-1}(x)),$$

où T^n est l'itérée n^{e} de T définie par récurrence par $\mathrm{T}^n = \mathrm{T} \circ \mathrm{T}^{n-1}$.

Préliminaire : le cas rationnel. Soient a, b deux entiers naturels tels que $0 < a < b$ et considérons le rationnel $x = a/b \in [0, 1[$. Montrer qu'il existe une unique suite finie $a_1, \ldots, a_n$ de nombres naturels $a_i \geqslant 1$ telle que

$$\frac{a}{b} = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}}.$$

On dit que l'on a développé a/b en fraction continue (finie). Pour simplifier la notation, nous noterons la fraction obtenue ci-dessus sous la forme

$$\frac{1}{a_1 +} \frac{1}{a_2 +} \cdots + \frac{1}{a_n}.$$

En déduire qu'il existe $n \geqslant 1$ tel que $a_n(x) = \infty$ si et seulement si $x \in \mathbf{Q} \cap [0, 1[$.

Dans toute la suite nous supposerons toujours que x est irrationnel. En particulier $a_n(x)$ est définie pour tout n et est un nombre entier et $\mathrm{T}^n(x) \in]0,1[$.

Partie I : identités et estimées de base. Rappelons que $x \in]0,1[$ est toujours dans ce qui suit supposé irrationnel.

1. Montrer par récurrence sur n que l'on a

$$x = \frac{1}{a_1(x)+}\frac{1}{a_2(x)+}\cdots+\frac{1}{a_n(x)+\mathrm{T}^n(x)}.$$

2. Montrer que pour tout $n \geqslant 1$ on a

$$p_{n-1}(x)q_n(x) - p_n(x)q_{n-1}(x) = (-1)^n,$$

où les $p_n(x), q_n(x)$ sont définies par les récurrence indiquées plus haut à partir de la suite $(a_n(x))$.

3. Montrer que pour tout $t \in [0,1]$, toute suite d'entiers naturels non nuls $(a_n)_{n\geqslant 1}$ et tout $n \geqslant 1$ on a l'identité

$$\frac{1}{a_1+}\cdots+\frac{1}{a_{n-1}+}\frac{1}{a_n+t} = \frac{p_n+tp_{n-1}}{q_n+tq_{n-1}}.$$

En déduire que l'on a

$$\frac{1}{a_1(x)+}\cdots+\frac{1}{a_n(x)} = \frac{p_n(x)}{q_n(x)}$$

et

$$x = \frac{p_n(x)+\mathrm{T}^n(x)p_{n-1}(x)}{q_n(x)+\mathrm{T}^n(x)q_{n-1}(x)}.$$

La fraction $p_n(x)/q_n(x)$ s'appelle le n^{e} quotient partiel de x.

4. Calculer la suite $(a_n(\sqrt{2}-1))_{n\geqslant 1}$.
5. Montrer que pour tout $n \geqslant 1$ on a

$$a_{n-1}(x) \leqslant \frac{1}{\mathrm{T}^n(x)} \leqslant a_{n+1}(x)+1$$

et en déduire que

$$\frac{1}{q_n(x)(q_n(x)+q_{n+1}(x))} \leqslant \left| x - \frac{p_n(x)}{q_n(x)} \right| \leqslant \frac{1}{q_n(x)q_{n+1}(x)}.$$

Que dire de la fraction $p_n(x)/q_n(x)$ parmi les nombres rationnels de dénominateur plus petit que $q_n(x)$?

6. Montrer que pour tout $n \geqslant 1$ on a $p_n(x) \geqslant 2^{(n-2)/2}$ et $q_n(x) \geqslant 2^{(n-1)/2}$ et en déduire que

$$\left| \mathrm{Log}\, \frac{x}{p_n(x)/q_n(x)} \right| \leqslant \frac{1}{2^{n-2}}.$$

Partie II : ergodicité de la transformation de Gauß. Pour toute suite $(a_n)_{n\geqslant 1}$ d'entiers naturels non nuls on note

$$\Delta_{a_1,\dots,a_n} = \{x \in [0,1[, a_1(x) = a_1, \dots, a_n(x) = a_n\}$$

et l'on définit, pour tout $t \in [0,1[$:

$$\psi_{a_1,\dots,a_n}(t) = \frac{p_n + tp_{n-1}}{q_n + tq_{n-1}},$$

où les p_n, q_n sont définies par les relations de récurrence mentionnées plus haut.

1. Montrer que $\Delta_{a_1,\dots,a_n} = \psi_{a_1,\dots,a_n}([0,1[)$ et en déduire que $\Delta_{a_1,\dots,a_n}$ est un intervalle qui satisfait à
$$|\Delta_{a_1,\dots,a_n}| = \frac{1}{q_n(q_n + q_{n-1})} \leqslant \frac{1}{2^{n-1}}.$$
2. Si A est une partie mesurable de $[0,1[$ on pose
$$\mu(\mathrm{A}) = \frac{1}{\mathrm{Log}\,2}\int_{\mathrm{A}} \frac{\mathrm{d}x}{1+x}.$$
Calculer $\mu([0,1[)$ et montrer que $\mu(\mathrm{A}) = 0$ si et seulement si $|\mathrm{A}| = 0$. On dit que μ est une mesure de probabilité à densité absolument continue sur $[0,1[$. On l'appelle la *mesure de Gauß.*
3. Soit $\alpha \in \,]0,1[$; montrer que
$$\mathrm{T}^{-1}([0,\alpha]) = \bigcup_{k=1}^{\infty} \left[\frac{1}{k+\alpha}, \frac{1}{k}\right]$$
et en déduire que
$$\int_0^\alpha \frac{\mathrm{d}x}{1+x} = \sum_{k=0}^{+\infty} \int_{(k+\alpha)^{-1}}^{k^{-1}} \frac{\mathrm{d}x}{1+x}.$$
4. Montrer que pour tout ensemble mesurable $\mathrm{A} \in [0,1[$ on a
$$\mu(\mathrm{T}^{-1}(\mathrm{A})) = \mu(\mathrm{A}).$$
Ainsi la transformation de Gauß préserve-t-elle la mesure de Gauß.
5. Si $\mathbf{P}$ est une probabilité sur $[0,1]$ (ici $\mathbf{P}$ sera soit la mesure de Lebesgue soit la mesure de Gauß) et si A, B sont deux parties mesurables de $[0,1[$ avec $\mathbf{P}(\mathrm{B}) > 0$, rappelons que l'on note $\mathrm{P}(\mathrm{A}|\mathrm{B})$ la probabilité de A conditionnellement à l'événement B, et que celle-ci vaut
$$\mathbf{P}(\mathrm{A}|\mathrm{B}) = \frac{\mathbf{P}(\mathrm{A}\cap\mathrm{B})}{\mathbf{P}(\mathrm{B})}.$$
Les deux événements A et B sont indépendants si le fait de savoir que B a lieu ne change pas la probabilité de A, si donc $\mathbf{P}(\mathrm{A}|\mathrm{B}) = \mathbf{P}(\mathrm{A})$.

Notons temporairement $\mathbf{L}$ la mesure de Lebesgue sur $[0,1[$ vue comme mesure de probabilité. On se donne une suite d'entiers $(a_n)_{n\geqslant 1}$ ainsi que deux réels x, y satisfaisant à $0 \leqslant x < y < 1$ et l'on note

$$\mathrm{I}_n = \big\{t \in [0,1[;\ x \leqslant \mathrm{T}^n(t) \leqslant y\big\} \cap \Delta_{a_1,\ldots,a_n}.$$

Montrer que l'on a

$$\mathbf{L}\big(\mathrm{T}^{-n}([x,y[)|\Delta_{a_1,\ldots,a_n}\big) = \frac{\psi_{a_1,\ldots,a_n}(y) - \psi_{a_1,\ldots,a_n}(x)}{\psi_{a_1,\ldots,a_n}(1) - \psi_{a_1,\ldots,a_n}(0)}$$

et en déduire que

$$\frac{1}{2}\big|[x,y[\big| \leqslant \mathbf{L}(\mathrm{T}^{-n}\big([x,y[\big)|\Delta_{a_1,\ldots,a_n}) \leqslant 2|[x,y[|.$$

6. Montrer que l'inégalité qui précède est encore vraie si l'on remplace $[x,y[$ par un ensemble mesurable quelconque $\mathrm{A} \subset [0,1[$. En déduire que pour tout ensemble mesurable $\mathrm{A} \subset [0,1[$ on a

$$\frac{\operatorname{Log} 2}{4}\mu(\mathrm{A}) \leqslant \mu\big(\mathrm{T}^{-n}(\mathrm{A})|\Delta_{a_1,\ldots,a_n}\big) \leqslant \frac{4}{\operatorname{Log} 2}\mu(\mathrm{A}).$$

7. Soit A un ensemble mesurable tel que $\mathrm{T}^{-1}(\mathrm{A}) = \mathrm{A}$ et $\mu(\mathrm{A}) > 0$. Montrer que

$$\frac{\operatorname{Log} 2}{4}\mu(\Delta_{a_1,\ldots,a_n}) \leqslant \mu(\Delta_{a_1,\ldots,a_n}|\mathrm{A})$$

et en déduire que pour tout ensemble $\mathrm{E} \in [0,1[$ mesurable on a

$$\frac{\operatorname{Log} 2}{4}\mu(\mathrm{E}) \leqslant \mu(\mathrm{E}|\mathrm{A}).$$

Montrer enfin que l'on a nécessairement $\mu(\mathrm{A}) = 1$; ainsi la mesure de Gauß est-elle une mesure invariante et ergodique pour la transformation de Gauß T. On peut donc appliquer le théorème ergodique de Birkhoff.

Partie III : propriétés statistiques du développement en fractions continues. Rappelons que comme μ est une mesure ergodique invariante pour la transformation de Gauß T et que comme $\mu(\mathrm{A}) = 0$ si et seulement si $|\mathrm{A}| = 0$, le théorème ergodique de Birkhoff affirme que pour toute fonction $\varphi \in \mathrm{L}^1([0,1[)$ on a, pour presque tout x :

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{k=0}^{n-1} \varphi(\mathrm{T}^k(x)) = \frac{1}{\operatorname{Log} 2}\int_0^1 \frac{\varphi(x)}{1+x}\,\mathrm{d}x.$$

Nous aurons dans cette partie à utiliser toutes les estimées et inégalités démontrées dans les deux parties précédentes. Pensez-à vous y référer !

1. Fixons un entier $k \geqslant 1$ et intéressons-nous à la fréquence asymptotique d'apparition du nombre k dans la suite des quotients partiels $a_1(x), a_2(x) \ldots$ du développement en fraction continue d'un nombre x. Montrer que celle-ci est égale à

$$\frac{1}{\operatorname{Log} 2} \operatorname{Log} \frac{(k+1)^2}{k(k+2)}.$$

On pourra avec profit considérer la fonction indicatrice de la partie $\{x;\ a_1(x) = k\}$. En déduire que la suite $(a_n(x))$ est non bornée pour presque tout x.

2. Montrer que pour tout $j \geqslant 1$ on a $p_{j+1}(x) = q_j(\mathrm{T}(x))$ et en déduire que

$$\frac{1}{q_n(x)} = \prod_{k=1}^{n} \frac{p_{n+1-k}(\mathrm{T}^{k-1}(x))}{q_{n+1-k}(\mathrm{T}^{k-1}(x))}.$$

3. Montrer que

$$\frac{1}{q_n(x)} = \prod_{k=1}^{n} \left(\frac{1}{a_k(x) + \cdots + \frac{1}{a_n(x)}} \right)$$

et en déduire que

$$\left| \operatorname{Log}(\mathrm{T}^{k-1}(x)) - \operatorname{Log}\left(\frac{1}{a_k(x) + \cdots + \frac{1}{a_n(x)}} \right) \right| \leqslant \frac{1}{2^{n-k-1}}$$

pour tout $1 \leqslant k \leqslant n$.

4. Montrer qu'il existe une constante $\mathrm{C} > 0$ telle que

$$\frac{1}{n} \operatorname{Log} q_n(x) = -\frac{1}{n} \sum_{k=0}^{n-1} \operatorname{Log}(\mathrm{T}^k(x)) + \frac{\mathrm{C}}{n}.$$

5. En utilisant ce qui précède et le théorème ergodique de Birkhoff, montrer que pour presque tout $x \in [0,1[$ on a

$$\lim_{n \to +\infty} \frac{1}{n} \operatorname{Log} q_n(x) = \frac{\pi^2}{12 \operatorname{Log} 2}.$$

On pourra utiliser sans démonstration l'identité

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)^2} = \frac{\pi^2}{12}.$$

6. Montrer que pour presque tout x on a

$$\lim_{n \to +\infty} \frac{1}{n} \operatorname{Log} \left| x - \frac{p_n(x)}{q_n(x)} \right| = -\frac{\pi^2}{6 \operatorname{Log} 2}.$$

This page intentionally left blank

Chapitre 13
Fascicule de résultats

Nous rassemblons ici les principaux résultats, définitions et propositions qu'il est indispensable de connaître. Nous les donnons systématiquement dans leur version la plus générale. Nous avons ensuite décrit une progression possible pour leur introduction et leur démonstration sous forme d'un graphe synthétique qui permettra à l'étudiant de s'y retrouver.

Fonctions d'une variable

Intégrabilité et intégrale

Définition 1 — *Soit* $\mathrm{I} = [a, b]$ *un intervalle compact. Une subdivision pointée* $\mathcal{P}$ *de* I *est la donnée d'une subdivision* $\mathrm{I}_1 = [x_0, x_1], \mathrm{I}_2 = [x_1, x_2], \ldots, \mathrm{I}_n = [x_{n-1}, x_n]$ *de* I *(avec* $a = x_0 \leqslant x_1 \leqslant \cdots \leqslant x_n = b$*) et d'un « pointage » de cette partition, c'est-à-dire de points* $t_1 \in \mathrm{I}_1, \ldots, t_n \in \mathrm{I}_n$*. On la note* $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$*, et on appelle les* t_i *des points de marquage de* $\mathcal{P}$*.*

Si I est un intervalle borné de bornes a, b, sa *longueur* est le nombre $|a - b|$ et elle est notée $|\mathrm{I}|$.

Définition 2 — *Soient* f *une fonction numérique quelconque définie sur l'intervalle* $[a, b]$*, et* $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$ *une subdivision pointée de* $[a, b]$*. On appelle somme de Riemann de* f *associée à* $\mathcal{P}$ *la quantité :*

$$\mathrm{S}(f, \mathcal{P}) = \sum_{k=1}^{n} |\mathrm{I}_k| f(t_k).$$

Définition 3 — *Un pas, ou* jauge *sur* $[a, b]$ *est une fonction* δ *définie sur* $[a, b]$ *et à valeurs dans* $]0, +\infty[$.

Définition 4 — *Une subdivision pointée* $\mathcal{P} = \{(\mathrm{I}_1, t_1), \ldots, (\mathrm{I}_n, t_n)\}$ *sera dite adaptée au pas, ou encore* δ*-fine, si, pour tout* $1 \leqslant k \leqslant n$ *on a*

$$\mathrm{I}_k \subseteq \left[t_k - \frac{\delta(t_k)}{2}, t_k + \frac{\delta(t_k)}{2}\right].$$

On a en particulier que $|\mathrm{I}_k| \leqslant \delta(t_k)$.

Lemme de Cousin — *Pour toute jauge* $\delta > 0$ *sur* $[a, b]$*, il existe une subdivision* δ*-fine.*

Définition 5 — *Soit* f *une fonction numérique définie sur* $[a, b]$*. La fonction* f *est dite intégrable sur* $[a, b]$ *s'il existe un nombre réel* S *tel que, pour tout* $\varepsilon > 0$*, il existe une jauge* δ_ε *sur* $[a, b]$ *telle que, pour toute subdivision* δ_ε*-fine* $\mathcal{P}$*, on ait*

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}| < \varepsilon.$$

Le nombre S *est appelé intégrale de la fonction* f *sur* $[a, b]$ *et est noté*

$$\int_{[a,b]} f(x)\,\mathrm{d}x \quad \textit{ou encore} \quad \int_a^b f(x)\,\mathrm{d}x, \quad \textit{voire} \quad \int_{[a,b]} f \quad \textit{ou} \quad \int_a^b f.$$

S'il existe, c'est l'unique nombre satisfaisant à la définition.

Définition 6 — *Une fonction* f *définie sur* $[a, b]$ *est dite* Riemann-intégrable *s'il existe* $\mathrm{S} \in \mathbf{R}$ *satisfaisant à la propriété suivante : pour tout* $\varepsilon > 0$*, il existe un* nombre réel $\delta > 0$ *tel que, pour toute partition* δ*-fine (pour la jauge* constante δ*), on ait*

$$|\mathrm{S} - \mathrm{S}(f, \mathcal{P})| < \varepsilon.$$

Définition 7 — *Une fonction* f *définie sur* $[a, b]$ *est dite* Lebesgue-intégrable *si* f *et* $|f|$ *sont intégrables sur* $[a, b]$*.*

Théorème 1 — *Toute fonction continue sur le segment* $[a, b]$ *y est à la fois intégrable, Riemann-intégrable et Lebesgue-intégrable.*

Théorème 2 — *Si* f *et* g *sont des fonctions intégrables sur* I*, et si* λ *est un nombre réel, alors* $f + g$ *et* λf *sont intégrables, et l'on a les identités*

$$\int_{\mathrm{I}} (f + g) = \int_{\mathrm{I}} f + \int_{\mathrm{I}} g \quad \textit{et} \quad \int_{\mathrm{I}} \lambda f = \lambda \int_{\mathrm{I}} f.$$

Théorème 3 — *Si* f *est intégrable sur* I *et vérifie* $f \geqslant 0$*, alors*

$$\int_{\mathrm{I}} f \geqslant 0.$$

Si f *et* g *sont intégrables sur* I *et vérifient* $f \geqslant g$*, alors*

$$\int_{\mathrm{I}} f \geqslant \int_{\mathrm{I}} g.$$

Théorème 4 (Relation de Chasles) — *Soient* $f : [a, b] \longrightarrow \mathbf{R}$ *et* $c \in]a, b[$. *Alors* f *est intégrable sur* $[a, b]$ *si, et seulement si, ses restrictions à* $[a, c]$ *et* $[c, b]$ *le sont. Dans ce cas, on a la relation*

$$\int_{[a,b]} f = \int_{[a,c]} f + \int_{[c,b]} f.$$

On note aussi l'intégrale de f *sur l'intervalle* $[a, b]$ *sous la forme*

$$\int_{[a,b]} f = \int_a^b f.$$

Intégrales et primitives

Formule fondamentale du calcul différentiel et intégral — *Soit* $\mathrm{F} : [a, b] \longrightarrow \mathbf{R}$ *dérivable sur* $[a, b]$, *de dérivée* f. *Alors* f *est intégrable et l'on a*

$$\int_a^b f = \mathrm{F}(b) - \mathrm{F}(a).$$

Théorème 5 — *Toute fonction continue* f *admet une primitive, égale à l'addition d'une constante près à la fonction* ψ_f *définie par*

$$\psi_f(x) = \int_a^x f, \;\; x \in [a, b].$$

Formule d'intégration par parties — *Soient* F *et* G *deux fonctions dérivables sur* $[a, b]$. *Alors* $\mathrm{F'G}$ *est intégrable si et seulement si* $\mathrm{FG'}$ *l'est. Dans ce cas, on a la formule*

$$\int_a^b \mathrm{F'G} = [\mathrm{FG}]_a^b - \int_a^b \mathrm{FG'}.$$

Formule du changement de variable — *Soient* g *une application dérivable sur l'intervalle compact* $[a, b]$, *et* f *une fonction admettant une primitive sur l'intervalle* $g([a, b])$. *Alors*

$$\int_{g(a)}^{g(b)} f(x)\,\mathrm{d}x = \int_a^b f(g(t))g'(t)\,\mathrm{d}t.$$

Intégrales impropres

Théorème 6 — *La fonction* $f : [a, +\infty[\longrightarrow \mathbf{R}$ *est intégrable si, et seulement si, pour tout* $c \in [a, +\infty[$, f *est intégrable sur* $[a, c]$, *et si la limite*

$$\lim_{c \to +\infty} \int_a^c f(x)\, \mathrm{d}x = \mathrm{L}$$

existe. On a alors

$$\int_a^{+\infty} f(x)\, \mathrm{d}x = \mathrm{L}.$$

Définition 8 — *Une fonction* f *définie sur un intervalle quelconque* I *y est dite localement intégrable si elle est intégrable sur tout intervalle compact inclus dans* I.

Définition 9 — *Soit* f *une fonction localement intégrable sur* $[a, +\infty[$. *L'intégrale*

$$\mathrm{I} = \int_a^{+\infty} f(x)\, \mathrm{d}x$$

est dite convergente *si* f *est intégrable sur* $[a, +\infty[$. *Elle est dite* absolument convergente *si* f *et* $|f|$ *sont intégrables sur* $[a, +\infty[$, *c'est-à-dire si* f *est Lebesgue-intégrable sur* $[a, +\infty[$. *On dit qu'elle est* semi-convergente *si elle est convergente mais pas absolument convergente. Si elle n'est pas convergente, on dit qu'elle est* divergente. *Le cas des autres types d'intervalle donne lieu à une terminologie analogue.*

Proposition 7 — *Soit* f *une fonction localement intégrable sur l'intervalle* $[a, +\infty[$.

1. *Si* $|f(x)| \leqslant x^{-\alpha}$ *pour un certain* $\alpha > 1$ *et pour tout* x *assez grand, alors* I *est absolument convergente.*
2. *Si* $|f(x)| \leqslant \exp(-\eta x)$ *pour un certain* $\eta > 0$ *et pour tout* x *assez grand, alors* I *est absolument convergente.*
3. *Si* $|f(x)| \geqslant x^{-\alpha}$ *pour un certain* $\alpha \leqslant 1$ *et pour tout* x *assez grand, alors* I *est divergente.*

Proposition 8 — *Soit* f *une fonction localement intégrable sur* $]a, b]$.

1. *Si* $|f(x)| \leqslant 1/(x-a)^\alpha$ *pour un certain* $\alpha < 1$ *et pour tout* x *assez proche de* a, *alors* I *est absolument convergente.*
2. *Si* $|f(x)| \leqslant |\eta \operatorname{Log}|x-a||$ *pour un certain* $\eta > 0$ *et pour tout* x *assez proche de* a, *alors* I *est absolument convergente.*
3. *Si* $|f(x)| \geqslant 1/(x-a)^\alpha$ *pour un certain* $\alpha \geqslant 1$ *et pour tout* x *assez proche de* a, *alors* I *est divergente.*

Théorème 9 — *Soit f une fonction* positive *sur l'intervalle $[a, +\infty[$. Pour que l'intégrale*

$$\mathrm{I} = \int_a^{+\infty} f(x)\,\mathrm{d}x$$

soit (absolument) convergente, il faut et il suffit qu'il existe une suite croissante (x_n), tendant vers $+\infty$, et telle que la série de terme général

$$u_n = \int_{x_n}^{x_{x+1}} f(x)\,\mathrm{d}x$$

soit (absolument) convergente.

Théorèmes de convergence

Définition 10 — *Une partie N de $\mathbf{R}$ est de mesure nulle si pour tout $\varepsilon > 0$, il existe une suite d'intervalles ouverts $(\mathrm{J}_k)_{k\in\mathbf{N}}$, telle que*

$$\mathrm{N} \subset \bigcup_{k\in\mathbf{N}} \mathrm{J}_k \quad \textit{et} \quad \sum_{k=0}^{+\infty} |\mathrm{J}_k| \leqslant \varepsilon.$$

Une propriété $\mathrm{P}(x)$ dépendant de x pour x dans une partie A de $\mathbf{R}$ est dite vraie presque partout *(ou* pour presque tout $x \in \mathrm{A}$*) si l'ensemble des x de A tels que $\mathrm{P}(x)$ soit violée est une partie de mesure nulle.*

Théorème 10 — *Soit f une fonction nulle presque partout sur $[a, b]$. Alors f est intégrable, d'intégrale nulle.*

Corollaire 11 — *Si f est intégrable sur $[a, b]$, et si g est une fonction sur $[a, b]$ égale à f presque partout, alors g est intégrable et*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x.$$

Théorème 12 — *Soient $a < b$ et f une fonction définie sur $[a, b]$, intégrable, positive, et d'intégrale nulle. Alors f est nulle presque partout.*

Notation : si I est un intervalle quelconque, nous noterons $\mathcal{L}(\mathrm{I})$ l'ensemble des fonctions intégrables au sens de Lebesgue sur I.

Proposition 13 — *L'espace $\mathcal{L}(\mathrm{I})$ est un espace vectoriel. De plus, $f \in \mathcal{L}(\mathrm{I})$ si et seulement si $f^+ = \max\{f, 0\}$ et $f_- = \min\{f, 0\}$ sont intégrables. Enfin, si $f, g \in \mathcal{L}(\mathrm{I})$, alors $\min\{f, g\}$ et $\max\{f, g\}$ sont aussi intégrables au sens de Lebesgue.*

Théorème de convergence monotone — *Soient* I *un intervalle quelconque,* (f_n) *une suite de fonctions définies sur* I *et* f *une fonction définie sur* I. *On suppose que*

1. *Les fonctions* f_n *sont Lebesgue-intégrables ;*
2. *Pour presque tout* $x \in \mathrm{I}$, *la suite* $(f_n(x))$ *est croissante ;*
3. *La suite* (f_n) *converge vers* f *presque partout, c'est-à-dire pour presque tout* $x \in \mathrm{I}$, *on a* $\lim_{n\to+\infty} f_n(x) = f(x)$ *;*
4. *Et la suite des intégrales* $\left(\int_\mathrm{I} f_n\right)$ *est majorée.*

Alors f *est Lebesgue-intégrable et*

$$\lim_{n\to+\infty} \int_\mathrm{I} f_n(x)\,\mathrm{d}x = \int_\mathrm{I} \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x = \int_\mathrm{I} f(x)\,\mathrm{d}x.$$

Le même énoncé est vrai en remplaçant « croissante » par « décroissante » et « majorée » par « minorée ».

Lemme de Fatou — *Soient* f_n, g *des fonctions Lebesgue-intégrables telles que* $g \leqslant f_n$ *pour tout* n *et* $\liminf \int_\mathrm{I} f_n$ *soit finie. Alors* $\liminf f_n$ *est Lebesgue-intégrable et*

$$(\dagger) \qquad -\infty < \int_\mathrm{I} g \leqslant \int_\mathrm{I} \liminf f_n \leqslant \liminf \int_\mathrm{I} f_n < +\infty.$$

Théorème de convergence dominée de Lebesgue — *Soit* (f_n) *une suite de fonctions Lebesgue-intégrables sur* I, *convergeant presque partout vers une fonction* f. *On suppose qu'il existe une fonction Lebesgue-intégrable* g *telle que*

$$|f_n(x)| \leqslant g(x), \quad \textit{pour presque tout } x \in \mathrm{I}, \ \textit{et} \ \ \forall n \in \mathbf{N}.$$

Alors f *est Lebesgue-intégrable sur* I, *et*

$$\lim_{n\to+\infty} \int_\mathrm{I} f_n = \int_\mathrm{I} \lim_{n\to+\infty} f_n = \int_\mathrm{I} f.$$

Théorème 14 (Passage d'une limite sous le signe intégral) — *En gardant les mêmes notations, soit* t_0 *un point de* T. *On suppose que :*

1. *Pour tout* $t \neq t_0$, *la fonction* $x \mapsto f(x,t)$ *est Lebesgue-intégrable en la variable* x *sur* I *;*
2. *Pour presque tout* $x \in \mathrm{I}$, *on a* $\lim_{t\to t_0} f(x,t) = f(x,t_0)$ *;*
3. *Il existe une fonction* g, *Lebesgue-intégrable sur* I, *telle que, pour tout* $t \neq t_0$ *et presque tout* $x \in \mathrm{I}$ *on ait*
$$|f(x,t)| \leqslant g(x).$$

Alors $x \mapsto f(x, t_0)$ *est Lebesgue-intégrable sur* I *et*

$$\lim_{t \to t_0} \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x = \int_{\mathrm{I}} f(x,t_0)\,\mathrm{d}x.$$

Théorème 15 (Continuité sous le signe intégral) — *Soient* $\mathrm{I} = [a,b]$ *et* $\mathrm{T} = [c,d]$*, et* $f(x,t)$ *une fonction définie sur* $[a,b] \times [c,d]$*. On suppose que :*

1. *Pour tout* $t \in \mathrm{T}$*, la fonction* $x \mapsto f(x,t)$ *est Lebesgue-intégrable en la variable* x *sur* I*,*
2. *pour tout* $x \in \mathrm{I}$*, la fonction* $t \mapsto f(x,t)$ *est continue en la variable* t *sur* T*,*
3. *il existe une fonction Lebesgue-intégrable* g *sur* I *telle que*

$$|f(x,t)| \leqslant g(x), \quad \forall (x,t) \in \mathrm{I} \times \mathrm{T}.$$

Alors la fonction

$$t \longmapsto \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x$$

est continue sur $\mathrm{T} = [c,d]$*.*

Théorème 16 (Dérivabilité sous le signe intégral) — *Avec les même notations que plus haut, on suppose que :*

1. *Pour tout* $t_0 \in \mathrm{T}$*, la fonction* $x \mapsto f(x,t_0)$ *est Lebesgue-intégrable sur* I *;*
2. *Pour tout* $x \in \mathrm{I}$*, la fonction* $t \mapsto f(x,t)$ *est dérivable, de dérivée*

$$f_t(x) = \frac{\partial f}{\partial t}(x,t)\,;$$

3. *Il existe* g*, Lebesgue-intégrable sur* I*, telle que*

$$|f_t(x)| \leqslant g(x),\ \forall x \in \mathrm{I} \textit{ et } \forall t \in \mathrm{T}.$$

Alors pour tout t*, les fonctions*

$$x \longmapsto f(x,t) \quad \textit{et} \quad x \longmapsto \frac{\partial f}{\partial t}(x,t)$$

sont Lebesgue-intégrables sur $[a,b]$*, et la fonction*

$$t \longmapsto \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x$$

est dérivable sur T*, de dérivée*

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathrm{I}} f(x,t)\,\mathrm{d}x = \int_{\mathrm{I}} \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x.$$

Théorème de dérivabilité de Lebesgue — *Soit f une fonction intégrable sur l'intervalle compact $[a,b]$. La fonction ψ_f définie par*

$$\psi_f(x) = \int_a^x f(t)\,\mathrm{d}t$$

est dérivable en presque tout point de $[a,b]$, de dérivée égale presque partout à $\psi_f'(x) = f(x)$.

Intégrale des fonctions de plusieurs variables

Notations : dans toute la suite, $d \geqslant 1$ est un entier fixé, la dimension de l'espace $\mathbf{R}^d$ sur lequel les fonctions que nous voudrons intégrer seront définies. Nous noterons en général un point de $\mathbf{R}^d$ en gras, sous la forme $\mathbf{x} = (x_1, \ldots, x_d)$. Si $1 \leqslant k \leqslant d$, on note π_k la projection sur la k^{e} coordonnée de sorte que $\pi_k(\mathbf{x}) = x_k$.

Intégrabilité et intégrale

Définition 11 — *Un pavé de $\mathbf{R}^d$ est une partie de $\mathbf{R}^d$ de la forme*

$$\mathrm{P} = \mathrm{I}_1 \times \mathrm{I}_2 \times \cdots \times \mathrm{I}_d,$$

où les I_k sont des intervalles compacts *de $\mathbf{R}$. Le volume (d-dimensionnel) de P est donné par le produit des longueurs des côtés*

$$|\mathrm{P}| = |\mathrm{I}_1| \cdots |\mathrm{I}_d|.$$

Définition 12 — *Soit P un pavé de $\mathbf{R}^d$. On appelle subdivision pointée de P une famille*

$$\mathcal{P} = \{(\mathrm{P}_1, \mathbf{t}_1), \ldots, (\mathrm{P}_n, \mathbf{t}_n)\}$$

telle que

1. *Pour tout i, P_i est un pavé de $\mathbf{R}^d$;*
2. *Pour tout i, $\mathbf{t}_i \in \mathrm{P}_i$ et $\mathrm{P}_i \subset \mathrm{P}$;*
3. *Les P_i recouvrent exactement P :*

$$\mathrm{P} = \bigcup_{i=1}^n \mathrm{P}_i\,;$$

4. *Les P_i sont d'intérieurs deux à deux disjoints : si $i \neq j$, alors $\mathring{\mathrm{P}}_i \cap \mathring{\mathrm{P}}_j = \emptyset$.*

Définition 13 — *Soient* P *un pavé,* $f : \mathrm{P} \longrightarrow \mathbf{R}$ *une fonction numérique définie sur* P *et* $\mathcal{P} = \{(\mathrm{P}_1, \mathbf{t}_1), \ldots, (\mathrm{P}_n, \mathbf{t}_n)\}$ *une subdivision pointée de* P*. La somme de Riemann de* f *associée à* P *est le nombre* $\mathrm{S}(f, \mathcal{P})$ *:*

$$\mathrm{S}(f, \mathcal{P}) = \sum_{i=1}^{n} f(\mathbf{t}_i)|\mathrm{P}_i|,$$

Définition 14 — *Une jauge sur un pavé* P *de* $\mathbf{R}^d$ *est une fonction* $\delta : \mathrm{P} \longrightarrow \mathbf{R}_+^*$*. On dit que la subdivision pointée*

$$\mathcal{P} = \{(\mathrm{P}_1, \mathbf{t}_1), \ldots, (\mathrm{P}_n, \mathbf{t}_n)\}$$

est δ*-fine si, pour tout* $i = 1, \ldots, n$ *et tout* $k = 1, \ldots, d$*, on a*

$$\pi_k(\mathrm{P}_i) \subset \left]\pi_k(\mathbf{t}_i) - \frac{\delta(\mathbf{t}_i)}{2}, \pi_k(\mathbf{t}_i) + \frac{\delta(\mathbf{t}_i)}{2}\right[.$$

Lemme de Cousin — *Pour tout pavé* P *et toute jauge* δ *sur* P *il existe une subdivision* δ*-fine de* P*.*

Définition 15 — *La fonction* f *est intégrable sur le pavé* P *si et seulement s'il existe un nombre* $\mathrm{S} \in \mathbf{R}$*, appelé* intégrale de f sur P, *satisfaisant à la condition suivante : pour tout* $\varepsilon > 0$*, il existe une jauge* δ *sur* P *telle que, pour toute subdivision* δ*-fine* $\mathcal{P}$ *de* P*, on ait*

$$|\mathrm{S}(f, \mathcal{P}) - \mathrm{S}| < \varepsilon.$$

On note l'intégrale de f *sur* P *sous les diverses formes suivantes :*

$$\int_{\mathrm{P}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad \textit{ou} \quad \int_{\mathrm{P}} f(x_1, \ldots, x_d)\,\mathrm{d}x_1 \ldots \mathrm{d}x_d \quad \textit{voire} \quad \int_{\mathrm{P}} f.$$

Définition 16 — *La fonction* $f : \mathrm{P} \longrightarrow \mathbf{R}$ *est dite intégrable au sens de Lebesgue, ou Lebesgue-intégrable, si* f et $|f|$ *sont intégrables sur* P*. On notera* $\mathcal{L}^1(\mathrm{P})$ *l'ensemble des fonctions Lebesgue-intégrables sur* P*.*

Définition 17 — *La fonction* $f : \mathbf{R}^d \longrightarrow \mathbf{R}$ *est dite Lebesgue-intégrable (sur* $\mathbf{R}^d$*) si pour tout entier* $n \geqslant 1$*,* f *est Lebesgue-intégrable sur* $\mathrm{P}_n = [-n, n]^d$*, et si les limites*

$$\lim_{n \to +\infty} \int_{\mathrm{P}_n} f(\mathbf{x})\,\mathrm{d}\mathbf{x} \quad \textit{et} \quad \lim_{n \to +\infty} \int_{\mathrm{P}_n} |f(\mathbf{x})|\,\mathrm{d}\mathbf{x}$$

existent et sont finies. On pose alors

$$\int_{\mathbf{R}^d} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \lim_{n \to +\infty} \int_{\mathrm{P}_n} f(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Définition 18 — *Soient* Ω *un ouvert de* $\mathbf{R}^d$ *et* $f : \Omega \longrightarrow \mathbf{R}$ *une fonction numérique définie sur* Ω. *Notons* $\bar{f}^\Omega$ *la fonction définie sur* $\mathbf{R}^d$ *obtenue en prolongeant* f *par* 0 *hors de* Ω. *On dit que* f *est Lebesgue-intégrable sur* Ω *si* $\bar{f}^\Omega$ *l'est sur* $\mathbf{R}^d$. *On note alors*

$$\int_\Omega f(\mathbf{x})\,\mathrm{d}\mathbf{x}$$

la valeur de son intégrale et $\mathcal{L}^1(\Omega)$ *l'espace des fonctions Lebesgue-intégrables sur* Ω.

Théorème 17 — *Soient* Ω *un ouvert* borné *de* $\mathbf{R}^d$, *et* f *une fonction définie et continue sur* l'adhérence $\bar{\Omega}$ *de* Ω. *Alors* f *est Lebesgue-intégrable sur* Ω.

Théorème 18 (Relation de Chasles) — *Soient* P *un pavé compact de* $\mathbf{R}^d$, $\mathcal{P} = \{\mathrm{P}_1, \ldots, \mathrm{P}_n\}$ *une subdivision de* P *en pavés et* $f : \mathrm{P} \longrightarrow \mathbf{R}$ *une fonction numérique. Alors* f *est Lebesgue-intégrable sur* P *si et seulement si* f *est Lebesgue-intégrable sur* P_i *pour tout* $i = 1, \ldots, n$. *Dans ce cas, on a de plus la formule*

$$\int_{\mathrm{P}} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \sum_{i=1}^{n} \int_{\mathrm{P}_i} f(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Théorèmes de convergence

Définition 19 — *Une partie* $\mathrm{N} \subset \mathbf{R}^d$ *est dite de mesure nulle si pour tout* $\varepsilon > 0$ *il existe une famille au plus dénombrable* $(\mathrm{P}_n)_{n\in\mathbf{N}}$ *de pavés de* $\mathbf{R}^d$ *tels que*

$$\mathrm{N} \subset \bigcup_{n\in\mathbf{N}} \mathrm{P}_n \quad \textit{et} \quad \sum_{i\in\mathbf{N}} |\mathrm{P}_n| < \varepsilon.$$

Une propriété $\mathrm{P}(\mathbf{x})$ *est dite vraie presque partout ou pour presque tout* $\mathbf{x}$ *si la propriété est vraie pour tout* $\mathbf{x}$ *hors d'un ensemble de mesure nulle.*

Théorème 19 — *Soient* Ω *un ouvert de* $\mathbf{R}^d$ *et* f *une fonction définie sur* Ω.

1. *Si* f *est nulle presque partout, alors* f *est Lebesgue-intégrable sur* Ω *et son intégrale est nulle.*
2. *Si* f *est positive et Lebesgue-intégrable, alors*

$$\int_\Omega f(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0$$

si et seulement si $f(\mathbf{x}) = 0$ *presque partout.*

Théorème de convergence monotone — *Soient* Ω *un ouvert de* $\mathbf{R}^d$, (f_n) *une suite de fonctions définies sur* Ω *et* f *une fonction définie sur* Ω. *On suppose que*

1. *Les fonctions* f_n *sont Lebesgue-intégrables ;*
2. *Pour presque tout* $\mathbf{x} \in \Omega$, *la suite* $(f_n(\mathbf{x}))$ *est croissante ;*
3. *La suite* (f_n) *converge presque partout vers* f, *c'est-à-dire pour presque tout* $\mathbf{x} \in \Omega$ *on a* $\lim_{n\to+\infty} f_n(\mathbf{x}) = f(x)$ *;*
4. *La suite de nombres réels* $\int_\Omega f_n(\mathbf{x})\,\mathrm{d}\mathbf{x}$ *est majorée.*

Alors f *est Lebesgue-intégrable et*

$$\lim_{n\to+\infty} \int_\Omega f_n(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_\Omega \lim_{n\to+\infty} f_n(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_\Omega f(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Le même énoncé est vrai en remplaçant « croissante » par « décroissante » et « majorée » par « minorée ».

Théorème de convergence dominée de Lebesgue — *Soit* (f_n) *une suite de fonctions Lebesgue-intégrables sur* Ω, *convergeant presque partout vers une fonction* f. *On suppose qu'il existe une fonction* g *Lebesgue-intégrable sur* Ω *telle que*

$$|f_n(\mathbf{x})| \leqslant g(\mathbf{x}), \quad \textit{pour presque tout}\, \mathbf{x} \in \Omega,\ \textit{et}\, \forall n \in \mathbf{N}.$$

Alors f *est Lebesgue-intégrable sur* Ω, *et*

$$\lim_{n\to+\infty} \int_{\mathrm{I}} f_n = \int_{\mathrm{I}} \lim_{n\to+\infty} f_n = \int_{\mathrm{I}} f.$$

Théorèmes spécifiques aux fonctions de plusieurs variables

Théorème et formule de Fubini — *Soient* P_1 *un pavé de* $\mathbf{R}^{d_1}$ *et* P_2 *un pavé de* $\mathbf{R}^{d_2}$ *avec* $d_1 + d_2 = d$. *Soient* $\mathrm{P} = \mathrm{P}_1 \times \mathrm{P}_2$ *et* f *une fonction définie sur* P : $f(\mathbf{x}) = f(\mathbf{x}_1, \mathbf{x}_2)$, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2) \in \mathrm{P} = \mathrm{P}_1 \times \mathrm{P}_2$. *Pour tout* $\mathbf{x}_1 \in \mathrm{P}_1$, *on définit la fonction partielle* $f(\mathbf{x}_1, \cdot) : \mathrm{P}_2 \longrightarrow \mathbf{R}$ *par* $f(\mathbf{x}_1, \cdot)(\mathbf{x}_2) = f(\mathbf{x}_1, \mathbf{x}_2)$. *On définit de même pour tout* $\mathbf{x}_2 \in \mathrm{P}_2$ *la fonction partielle* $f(\cdot, \mathbf{x}_2)$. *Si* f *est Lebesgue-intégrable, alors*

1. *Pour presque tout* $\mathbf{x}_1 \in \mathrm{P}_1$, *la fonction* $f(\mathbf{x}_1, \cdot) : \mathrm{P}_2 \longrightarrow \mathbf{R}$ *est Lebesgue-intégrable sur* P_2 *;*
2. *Pour presque tout* $\mathbf{x}_2 \in \mathrm{P}_2$, *la fonction* $f(\cdot, \mathbf{x}_2) : \mathrm{P}_1 \longrightarrow \mathbf{R}$ *est Lebesgue-intégrable sur* P_1 *;*
3. *Pour presque tout* $\mathbf{x}_1 \in \mathrm{P}_1$, *la fonction*

$$\mathbf{x}_1 \longmapsto \int_{\mathrm{P}_2} f(\mathbf{x}_1, \mathbf{x}_2)\,\mathrm{d}\mathbf{x}_2$$

est Lebesgue-intégrable sur P_2 *;*

4. Pour presque tout $\mathbf{x}_2 \in \mathrm{P}_2$*, la fonction*

$$\mathbf{x}_2 \longmapsto \int_{\mathrm{P}_1} f(\mathbf{x}_1, \mathbf{x}_2)\,\mathrm{d}\mathbf{x}_2$$

est Lebesgue-intégrable sur P_1*.*

On a de plus la formule suivante, dite « formule de Fubini » :

$$\int_{\mathrm{P}_1 \times \mathrm{P}_2} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathrm{P}_1} \left(\int_{\mathrm{P}_2} f(\mathbf{x}_1, \mathbf{x}_2)\,\mathrm{d}\mathbf{x}_2 \right) \mathrm{d}\mathbf{x}_1 = \int_{\mathrm{P}_2} \left(\int_{\mathrm{P}_1} f(\mathbf{x}_1, \mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1 \right) \mathrm{d}\mathbf{x}_2.$$

Théorème de Tonelli — *Soit* $f : \mathrm{P} = \mathrm{P}_1 \times \mathrm{P}_2 \longrightarrow \mathbf{R}$ *une fonction* positive. *Si l'une des trois intégrales*

$$\int_{\mathrm{P}} f(\mathbf{x})\,\mathrm{d}\mathbf{x}, \int_{\mathrm{P}_1} \left(\int_{\mathrm{P}_2} f(\mathbf{x}_1, \mathbf{x}_2)\,\mathrm{d}\mathbf{x}_2 \right) \mathrm{d}\mathbf{x}_1 \int_{\mathrm{P}_2} \left(\int_{\mathrm{P}_1} f(\mathbf{x}_1, \mathbf{x}_2)\,\mathrm{d}\mathbf{x}_1 \right) \mathrm{d}\mathbf{x}_2$$

existe (et est finie), alors les deux autres aussi, et les trois sont égales. En particulier, f *est Lebesgue-intégrable sur* P*.*

Définition 20 — *Soient* Ω *un ouvert de* $\mathbf{R}^d$ *et* $\varphi : \Omega \longrightarrow \mathbf{R}^d$ *une application différentiable. Le jacobien de* φ *en un point* $\mathbf{x}$ *de* Ω *est le déterminant de la différentielle de* φ *au point* $\mathbf{x}$ *:*

$$(\mathrm{Jac}\,\varphi)(\mathbf{x}) = \det \mathrm{D}_{\mathbf{x}}\varphi = \begin{vmatrix} \frac{\partial \varphi_1}{\partial x_1}(\mathbf{x}) & \frac{\partial \varphi_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial \varphi_1}{\partial x_d}(\mathbf{x}) \\ \vdots & & & \vdots \\ \frac{\partial \varphi_d}{\partial x_1}(\mathbf{x}) & \frac{\partial \varphi_d}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial \varphi_d}{\partial x_d}(\mathbf{x}) \end{vmatrix},$$

où les φ_i *sont les composantes de* φ*.*

Formule du changement de variable ou du jacobien — *Soient* Ω *un ouvert de* $\mathbf{R}^d$*,* $\varphi : \Omega \longrightarrow \varphi(\Omega) \subset \mathbf{R}^d$ *un difféomorphisme de classe* $\mathcal{C}^1$ *et* $f : \varphi(\Omega) \longrightarrow \mathbf{R}$ *une fonction Lebesgue-intégrable. Alors* $f \circ \varphi$ *est Lebesgue-intégrable sur* Ω *et on a la formule*

$$\int_{\varphi(\Omega)} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\Omega} f \circ \varphi(\mathbf{u}) |(\mathrm{Jac}\,\varphi)(\mathbf{u})|\,\mathrm{d}\mathbf{u}.$$

Formule de Green-Riemann — *Soient* K *un compact à bord régulier de* $\mathbf{R}^2$*,* $\alpha = \mathrm{P}\,\mathrm{d}x + \mathrm{Q}\,\mathrm{d}y$ *une* 1*-forme différentielle définie sur un voisinage de* K*. Alors*

$$\int_{\partial \mathrm{K}} (\mathrm{P}\,\mathrm{d}x + \mathrm{Q}\,\mathrm{d}y) = \iint_{\mathrm{K}} \left(\frac{\partial \mathrm{Q}}{\partial x} - \frac{\partial \mathrm{P}}{\partial y} \right) \mathrm{d}x\,\mathrm{d}y.$$

Fonctions mesurables et espaces de Lebesgue

Définition 21 — *Soient Ω un ouvert de $\mathbf{R}^d$ et $f : \Omega \longrightarrow \mathbf{R}$ une fonction numérique. On dit que f est mesurable (sur Ω) s'il existe une suite $(f_k)_{k\in\mathbf{N}}$ de fonctions continues sur Ω telle que pour presque tout $\mathbf{x} \in \Omega$ on ait $f(\mathbf{x}) = \lim_{k\to+\infty} f_k(\mathbf{x})$.*

Théorème 20 — *Si $f : \Omega \longrightarrow \mathbf{R}$ est mesurable et s'il existe une fonction g Lebesgue-intégrable sur Ω telle que $|f(\mathbf{x})| \leqslant g(\mathbf{x})$ presque partout, alors f est Lebesgue-intégrable sur Ω. Si $f : \Omega \longrightarrow \mathbf{R}$ est Lebesgue-intégrable sur $\mathbf{R}^d$, alors f est mesurable.*

Théorème 21 — *Soient $f, g, h \in \mathcal{M}(\Omega)$ ainsi que $\lambda \in \mathbf{R}$.*

1. *Les fonctions λf, $f + g$, $f - g$ et fg sont dans $\mathcal{M}(\Omega)$;*
2. *Si $f(\mathbf{x}) \neq 0$ presque partout, alors $1/f \in \mathcal{M}(\Omega)$;*
3. *Si $k(\mathbf{x}) = f(\mathbf{x})$ presque partout, alors $k \in \mathcal{M}(\Omega)$;*
4. *Si $\varphi : \mathbf{R} \longrightarrow \mathbf{R}$ est continue, alors $\varphi \circ f \in \mathcal{M}(\mathbf{R}^d)$;*
5. *La fonction f est mesurable si et seulement si f^+ et f^- le sont ;*
6. *Les fonctions f et g sont mesurables si et seulement si $\max\{f, g\}$ et $\min\{f, g\}$ le sont ;*
7. *Si f, g, h sont mesurables, alors $\mathrm{méd}\{f, g, h\}$ l'est.*

Définition 22 — *Si f est Lebesgue-intégrable sur l'ouvert Ω, sa semi-norme L^1 est par définition*

$$\|f\|_{\mathrm{L}^1} = \int_\Omega |f(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

On dit qu'une suite (f_n) de fonctions Lebesgue-intégrables converge en moyenne vers la fonction f si

$$\lim_{n\to+\infty} \|f - f_n\|_{\mathrm{L}^1} = 0.$$

Théorème 22 — *Soit $\mathcal{N}(\Omega)$ l'ensemble des fonctions nulles presque partout dans Ω. La semi-norme L^1 munit l'espace vectoriel quotient $\mathrm{L}^1(\Omega) = \mathcal{L}^1(\Omega)/\mathcal{N}(\Omega)$ d'une norme qui fait de $\mathrm{L}^1(\Omega)$ un espace de Banach.*

Théorème de Beppo Levi — *Soit (ψ_k) une suite d'éléments de l'espace $\mathcal{L}^1(\mathbf{R}^d)$ telle que la série*

$$\sum_{k=0}^{+\infty} \|\psi_k\|_{\mathrm{L}^1}$$

soit convergente, de somme K. Alors la série

$$\sum_{k=0}^{+\infty} \psi_k(\mathbf{x})$$

converge absolument presque partout, la fonction somme est une fonction Lebesgue-intégrable et

$$\int_{\mathbf{R}^d} \left(\sum_{k=0}^{+\infty} \psi_k(\mathbf{x}) \right) \mathrm{d}\mathbf{x} = \sum_{k=0}^{+\infty} \int_{\mathbf{R}^d} \psi_k(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

Proposition 23 — *Si une suite* (f_n) *d'éléments de* $\mathcal{L}^1(\mathbf{R}^d)$ *converge en moyenne vers* $f \in \mathcal{L}^1(\mathbf{R}^d)$*, alors il existe une sous-suite* (f_{n_k}) *qui converge presque partout vers* f*.*

Définition 23 — *Soient* $p \geqslant 1$ *un nombre réel et* Ω *un ouvert de* $\mathbf{R}^d$*. L'ensemble* $\mathcal{L}^p(\Omega)$ *des fonctions mesurables sur* Ω *et dont la puissance* p*-ème est Lebesgue-intégrable est un espace vectoriel. La semi-norme* L^p *de* $f \in \mathcal{L}^p(\Omega)$ *est, par définition :*

$$\|f\|_{\mathrm{L}^p} = \left(\int_\Omega |f(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} \right)^{1/p}.$$

Théorème 24 — *Soit* $\mathcal{N}(\Omega)$ *l'ensemble des fonctions nulles presque partout dans* Ω*. La semi-norme* L^p *munit l'espace vectoriel quotient* $\mathrm{L}^p(\Omega) = \mathcal{L}^p(\Omega)/\mathcal{N}(\Omega)$ *d'une norme qui fait de* $\mathrm{L}^p(\Omega)$ *un espace de Banach.*

Théorème 25 — *Soient* Ω *un ouvert de* $\mathbf{R}^d$ *et* f, g *deux fonctions mesurables définies sur* Ω*.*

1. **Inégalité de Cauchy-Schwarz :** *si* f *et* g *sont de carré intégrable sur* Ω*, alors* fg *est Lebesgue-intégrable sur* Ω *et*

$$\int_\Omega |f(\mathbf{x})g(\mathbf{x})| \, \mathrm{d}\mathbf{x} \leqslant \left(\int_\Omega f(\mathbf{x})^2 \, \mathrm{d}\mathbf{x} \right)^{1/2} \left(\int_\Omega g(\mathbf{x})^2 \, \mathrm{d}\mathbf{x} \right)^{1/2}.$$

2. **Inégalité de Hölder :** *soit* $p > 1$*, et* $q > 1$ *tels que* $1/p + 1/q = 1$*. Si* $f \in \mathrm{L}^p(\Omega)$ *et* $g \in \mathrm{L}^q(\Omega)$*, alors* $fg \in \mathrm{L}^1(\Omega)$ *et*

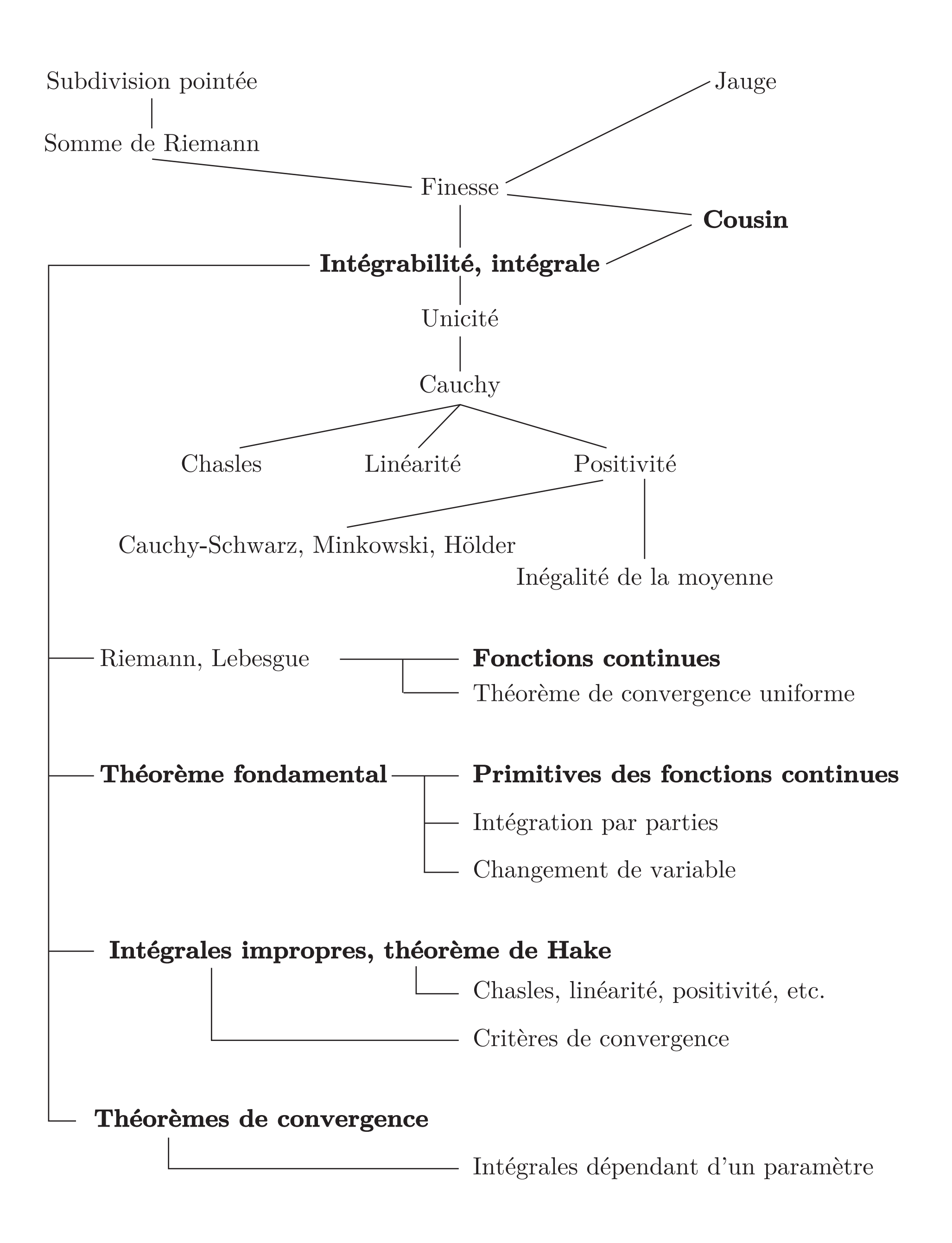
$$\int_\Omega |f(\mathbf{x})g(\mathbf{x})| \, \mathrm{d}\mathbf{x} \leqslant \left(\int_\Omega f(\mathbf{x})^p \, \mathrm{d}\mathbf{x} \right)^{1/p} \left(\int_\Omega g(\mathbf{x})^q \, \mathrm{d}\mathbf{x} \right)^{1/q}.$$

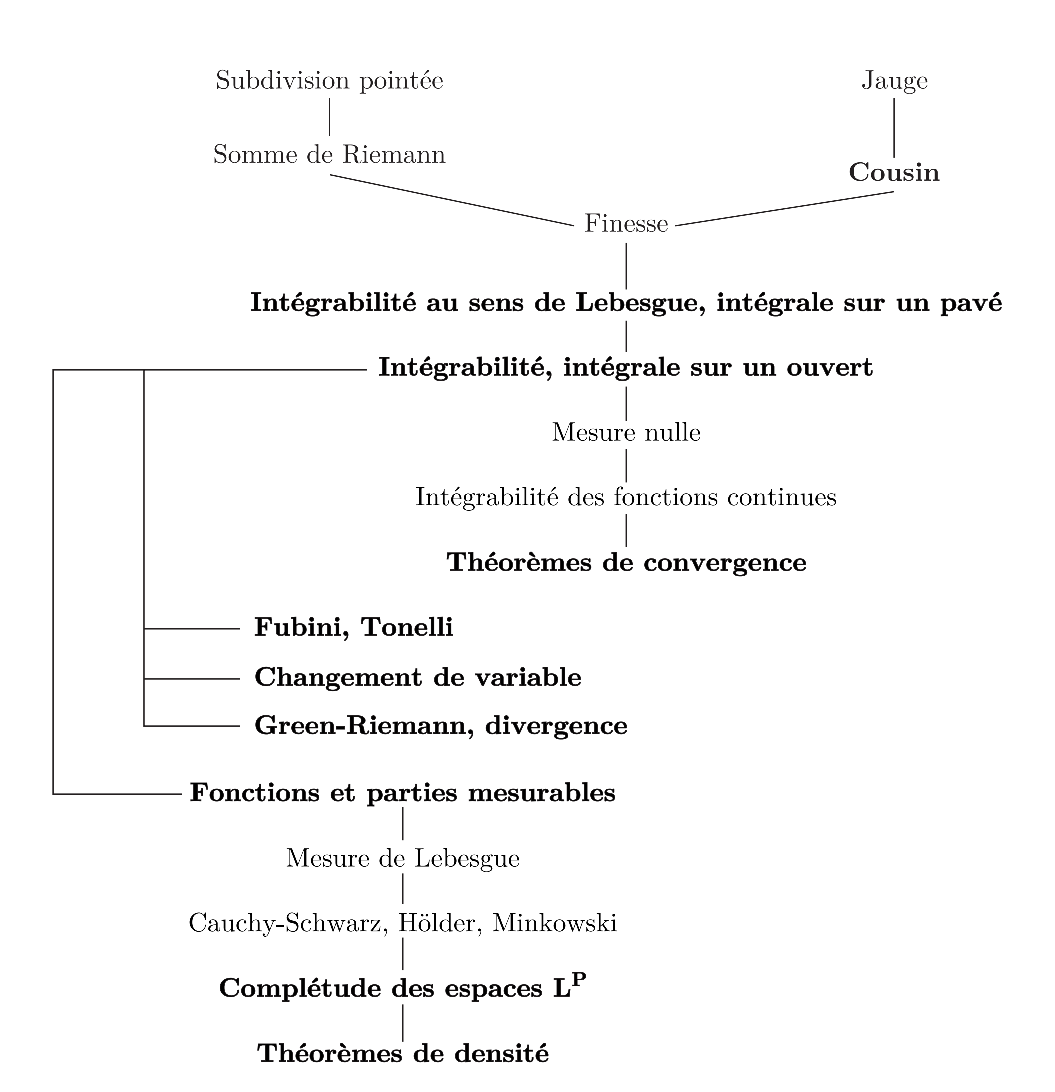
3. **Inégalité de Minkowski :** *soit* $p \geqslant 1$*. Si* $f, g \in \mathrm{L}^p(\Omega)$ *alors* $f + g \in \mathrm{L}^p(\Omega)$ *et*

$$\left(\int_\Omega |f(\mathbf{x}) + g(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} \right)^{1/p} \leqslant \left(\int_\Omega |f(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} \right)^{1/p} + \left(\int_\Omega |g(\mathbf{x})|^p \, \mathrm{d}\mathbf{x} \right)^{1/p}.$$

Théorème 26 — *Soient* Ω *un ouvert de* $\mathbf{R}^d$*,* $p \in [1, +\infty[$ *et* $k \in \mathbf{N} \cup \{\infty\}$*. Toute fonction* $f \in \mathrm{L}^p(\Omega)$ *est limite en norme* L^p *et presque partout d'une suite de fonctions de classe* $\mathcal{C}^k$ *à supports compacts dans* Ω*.*

Théorème 27 — *Une fonction en escalier sur un ouvert* Ω *de* $\mathbf{R}^d$ *est une combinaison linéaire (en particulier une somme finie) de fonctions indicatrices de pavés compacts inclus dans* Ω*. Soient* $p \geqslant 1$ *et* f *un élément de* $\mathrm{L}^p(\mathbf{R}^d)$*. Alors il existe une suite* $(h_n)_{n\in\mathbf{N}}$ *de fonctions en escalier intégrables sur* Ω *qui converge en norme* L^p *et presque partout vers* f*.*


Subdivision pointée
Somme de Riemann
Jauge
Finesse
Cousin
Intégrabilité, intégrale
Unicité
Cauchy
Chasles
Linéarité
Positivité
Cauchy-Schwarz, Minkowski, Hölder
Inégalité de la moyenne
Riemann, Lebesgue
Fonctions continues
Théorème de convergence uniforme
Théorème fondamental
Primitives des fonctions continues
Intégration par parties
Changement de variable
Intégrales impropres, théorème de Hake
Chasles, linéarité, positivité, etc.
Critères de convergence
Théorèmes de convergence
Intégrales dépendant d'un paramètre


Subdivision pointée
Jauge
Somme de Riemann
Cousin
Finesse
Intégrabilité au sens de Lebesgue, intégrale sur un pavé
Intégrabilité, intégrale sur un ouvert
Mesure nulle
Intégrabilité des fonctions continues
Théorèmes de convergence
Fubini, Tonelli
Changement de variable
Green-Riemann, divergence
Fonctions et parties mesurables
Mesure de Lebesgue
Cauchy-Schwarz, Hölder, Minkowski
Complétude des espaces L^P
Théorèmes de densité

This page intentionally left blank

Bibliographie

Nous avons regroupé ci-après quelques références bibliographiques. Il existe une multitude de livres et d'articles au sujet de l'intégration et nous nous sommes limités à quelques articles historiques ainsi qu'à quelques livres traitant de l'intégrale de Kurzweil et Henstock, ce sujet étant plus rarement évoqué.

Nous avons tout d'abord inclus les jalons historiques du sujet : le mémoire de Bernhard Riemann où il définit l'intégrale portant son nom, la note de Henri Lebesgue de 1901, les articles d'Arnaud Denjoy et Oskar Perron où est construite une intégrale équivalente à celle de Kurzweil et Henstock et enfin les deux articles fondateurs de ces derniers auteurs.

Nous avons ensuite cité les références de quelques textes où il est question de l'intégrale de jauge. Un seul, celui de Jean Mawhin, est en français. Celui de Robert Bartle est très touffu et va fort loin, tandis que ceux de Lee Peng Yeng et Rudolf Výborný ou de Robert McLeod sont plus accessibles et conformes au niveau du présent traité. Jean-Pierre Demailly a écrit un excellent cours d'intégration utilisant l'intégrale de jauge, mais il n'est pas édité et n'est disponible que sur sa page ouaibe.

Il existe par ailleurs de très nombreux cours d'analyse pour la licence, les classes préparatoires, les masters ou la préparation à l'agrégation et au CAPES. Les bibliothèques universitaires en proposent un grand nombre. On trouve, sur le marché de l'occasion, d'excellents livres destinés aux premiers cycles universitaires ou aux classes préparatoires. C'est un moyen peu onéreux d'acquérir des livres de mathématiques qui, malgré leur âge, gardent toute leur valeur.

[1] Bernhard Riemann. *Über die Darstellbarkeit einer Function durch eine trigonometrische Reihe.* Thèse d'habilitation, 1854.

[2] Henri Lebesgue. *Sur une généralisation de l'intégrale définie.* Comptes rendus des séances de l'Académie des sciences, **132**, p. 1025–1027, 1901.

[3] Arnaud Denjoy. *Une extension de l'intégrale de M. Lebesgue.* Comptes rendus des séances de l'Académie des sciences, **154**, p. 859–862, 1912.

[4] Oskar Perron. *Über den integralbegriff.* Sitzber. Heidelberg Akad. Wiss., Math.-Naturw. Klasse Abt. A, **16**, p. 1–16, 1914.

[5] Ralph Henstock. *The efficiency of convergence factors for functions of a continuous real variable.* J. London Math. Soc., **30**, p. 273–286, 1955.

[6] Jaroslaw Kurzweil. *Generalized ordinary differential equations and continuous dependance on a parameter.* Czech. Math. J., **7 (82)**, p. 418–446, 1957.

[7] Jean Mawhin. *Analyse. Fondements, Techniques, Évolutions* (De Broeck Université, Bruxelles), deuxième édition, 1997.

[8] Robert G. Bartle. *A Modern Theory Of Integration*, tome 32 de *Graduate Studies in Mathematics* (American Mathematical Society, Providence, RI), 2001.

[9] Robert M. McLeod. *The Generalized Riemann Integral*, tome 20 de *Carus Monograph* (Mathematical Association Of America, Washington, DC), 1980.

[10] Lee Peng Yeng et Rudolf Výborný. *The Integral. An Easy Approach after Kurzweil And Henstock* (Cambridge University Press, Cambridge, UK), 2000.

Index

A

B

C

D

E

F

G

H

I

J

K

L

M

N

P

R

S

T

U

V

